

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 08:03:07 ; Search time 7429 Seconds

(without alignments)
6849,471 Million cell updates/sec

Title: US-10-007-521-11

Sequence: 1174

1 GAGCGACGACCCCTCAAGCTG.....AAAAAAAAAAAAAAAAAAAA 1174

Scoring table:

IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 segs, 21671516395 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*
41: gb_ph:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1174	100.0	1174	6 A92311	A92311 Sequence 1
2	1174	100.0	1174	6 AR075389	AR075389 Sequence 1
3	1174	100.0	1174	6 AR094310	AR094310 Sequence 1
4	451.6	38.5	922	6 AR68070	AR68070 Sequence 65
5	451.6	38.5	922	6 AR163167	AR163167 Sequence 65
6	437.4	37.3	1060	6 A21793	A21793 H. insolens
7	437.4	37.3	1060	6 A23635	A23635 H. insolens
8	437.4	37.3	1060	6 A23644	A23644 H. insolens
9	437.4	37.3	1060	6 A23953	A23953 Endoglucanase
10	437.4	37.3	1060	6 A23957	A23957 Endoglucanase
11	437.4	37.3	1060	6 A41658	A41658 Sequence 1
12	437.4	37.3	1060	6 A68060	A68060 Sequence 55
13	437.4	37.3	1060	6 AR059002	AR059002 Sequence 1
14	437.4	37.3	1060	6 AR072921	AR072921 Sequence 1
15	437.4	37.3	1060	6 AR163160	AR163160 Sequence 1
16	437.4	37.3	1060	6 A13884	A13884 Sequence 1
17	437.4	37.3	1060	6 A13884	A13884 Sequence 1
18	437.4	37.3	1060	6 A13884	A13884 Sequence 1
19	437.4	37.3	1060	6 A13884	A13884 Sequence 1
20	437.4	37.3	1060	6 A13884	A13884 Sequence 1
21	437.4	37.3	1060	6 A13884	A13884 Sequence 1
22	437.4	37.3	1060	6 A13884	A13884 Sequence 1
23	429.6	35.0	919	6 A68072	A68072 Sequence 67
24	411.4	35.0	919	6 A68072	A68072 Sequence 67
25	411.4	35.0	919	6 A68072	A68072 Sequence 67
26	393.4	33.5	155568	8 NCBI9A17	BX284763 Neurospor
27	392.2	33.4	807	6 AR261284	AR261284 Sequence 73
28	392	33.4	922	6 AR163171	AR163171 Sequence 1
29	377.2	32.1	1257	6 AR121354	AR121354 Sequence 1
30	377.2	32.1	1257	6 AR121354	AR121354 Sequence 1
31	377.2	32.1	1257	6 AR121354	AR121354 Sequence 1
32	377.2	32.1	1257	6 AR121354	AR121354 Sequence 1
33	377.2	32.1	1257	6 AR121354	AR121354 Sequence 1
34	375.6	32.0	1257	6 AR213263	AR213263 Sequence 1
35	374.6	31.9	1904	6 AR003107	AR003107 Humicola
36	369.8	31.5	1154	6 AR094308	AR094308 Sequence 1
37	364.6	31.1	924	6 A68074	A68074 Sequence 69
38	364.6	31.1	924	6 A68074	A68074 Sequence 69
39	353.6	30.1	936	8 MAL515703	AJ515703 Melanocar
40	337	28.7	894	6 AR094306	AR094306 Sequence 1
41	336.2	28.6	912	6 A68076	A68076 Sequence 71
42	336.2	28.6	912	6 AR163170	AR163170 Sequence 1
43	334.8	28.5	927	6 AR094307	AR094307 Sequence 1
44	315.2	26.8	1443	8 FSOXCH	L29381 Fusarium ox
45	315.2	26.8	1473	6 A21795	A21795 F. oxysporum

ALIGNMENTS

RESULT 1
LOCUS A92311 1174 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent EP0843041.
ACCESSION A92311
VERSION A92311.1 GI:6741083
KEYWORDS
SOURCE Saccharomyces cerevisiae
ORGANISM Saccharomyces cerevisiae (baker's yeast)
REFERENCE 1 (bases 1 to 1174)
AUTHORS Kalum, L. and Lund, H.
TITLE Garments with considerable variation in abrasion level and process for its production using cellulolytic enzymes

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: EP 0843041-A 1 20-MAY-1998;
NOVONORDISK AS (DK)
FEATURES Location/Qualifiers
source 1..1174
/organism="Saccharomyces cerevisiae"
/mol_type="unassigned DNA"
/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 1174; DB 6; Length 1174;
Best Local Similarity 100.0%; Pred. No. 3.2e-199;
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAGGAGCAGCCCTCAAGCTGACAGTTTCACCCCGCTCTTTTTCGAGCCCGCAGGA 60
1 GAGGAGCAGCCCTCAAGCTGACAGTTTCACCCCGCTCTTTTTCGAGCCCGCAGGA 60
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGAGCTCCG 120
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGAGCTCCG 120
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATCTGAGCACTGGCAAGCCGTCGCTT 180
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATCTGAGCACTGGCAAGCCGTCGCTT 180
181 GAGCCGGGAAAGCCGCGCTCAGCCACCGGCTCAAGCTGCAATGCACTTCCAGGCC 240
181 GAGCCGGGAAAGCCGCGCTCAGCCACCGGCTCAAGCTGCAATGCACTTCCAGGCC 240
241 TGTCCGACTTCATATGCCAGTGGCGGCTGGCAAGCGAGCTGGCCGCTCAGCCAGC 300
241 TGTCCGACTTCATATGCCAGTGGCGGCTGGCAAGCGAGCTGGCCGCTCAGCCAGC 300
301 AGACTCCCTGGCGGCTGAGACGACATCTGCCTCAAGCGCTTCCGCGGAGAGATCCG 360
301 AGACTCCCTGGCGGCTGAGACGACATCTGCCTCAAGCGCTTCCGCGGAGAGATCCG 360
361 GCGGCTCCGAAATCTCTGAGTGGCTGGCGCTGCAAGCCGCTCACTTCGAGTCCG 420
361 GCGGCTCCGAAATCTCTGAGTGGCTGGCGCTGCAAGCCGCTCACTTCGAGTCCG 420
421 TGGCGGCAAGACATATGTTGAGTCAAGCAAGCACTGGCGGAGCACTGGAGATTAAC 480
421 TGGCGGCAAGACATATGTTGAGTCAAGCAAGCACTGGCGGAGCACTGGAGATTAAC 480
481 AGTTGATATGCAATGCGCGCGCGCGCTGGAGCATTTCAAGCGCTGAGCTGCGAGT 540
481 AGTTGATATGCAATGCGCGCGCGCGCTGGAGCATTTCAAGCGCTGAGCTGCGAGT 540
541 TGGGCGGCTTCCCGGCGCTCAATAGCGAGGCAATTTGCTGGCGGCAAGAGTGGT 600
541 TGGGCGGCTTCCCGGCGCTCAATAGCGAGGCAATTTGCTGGCGGCAAGAGTGGT 600
601 TCCCGGCGCGCTCAAGCCGCGCTGCAAGTGGCGGCTTGAATGCTTCAAGAGCGCGCA 660
601 TCCCGGCGCGCTCAAGCCGCGCTGCAAGTGGCGGCTTGAATGCTTCAAGAGCGCGCA 660
661 ACCCGAGCTTCAAGCTTCAAGCGAGTGGCGGCTGCGCGCGAGATGTTGCGCGCTCG 720
661 ACCCGAGCTTCAAGCTTCAAGCGAGTGGCGGCTGCGCGCGAGATGTTGCGCGCTCG 720
721 GCAAGCGCAAGAGCACTCAAGCTTCCCGCTTCAAGCCCGCAAGCGGAGGAGAGG 780
721 GCAAGCGCAAGAGCACTCAAGCTTCCCGCTTCAAGCCCGCAAGCGGAGGAGAGG 780
781 GCAAGCGCAAGAGCACTCAAGCTTCCCGCTTCAAGCCCGCAAGCGGAGGAGAGG 840
781 GCAAGCGCAAGAGCACTCAAGCTTCCCGCTTCAAGCCCGCAAGCGGAGGAGAGG 840
841 GTGGCTGCAAGCTTCAAGAGTGGGCTCAAGTGGGCTCAAGTGGGCTTCAAGCGAGTCA 900
841 GTGGCTGCAAGCTTCAAGAGTGGGCTCAAGTGGGCTCAAGTGGGCTTCAAGCGAGTCA 900

901 CTTGCTCTTGGGACCAAGCTTCCGCAAGAGTTGAAGCACTAATCTCCAGTGGCTTAA 960
901 CTTGCTCTTGGGACCAAGCTTCCGCAAGAGTTGAAGCACTAATCTCCAGTGGCTTAA 960
961 CAGCTTTTGGCAAGAGTGGCGGAGCAAGCAAGAGCCGTCATCTTCTGATGATAT 1020
961 CAGCTTTTGGCAAGAGTGGCGGAGCAAGCAAGAGCCGTCATCTTCTGATGATAT 1020
1021 TTTTGAAGCGTCAATCATATTAATCTTGAATCTTGTATATGACAGCGCGTACACA 1080
1021 TTTTGAAGCGTCAATCATATTAATCTTGAATCTTGTATATGACAGCGCGTACACA 1080
1081 TCTCAGACCGCACTTTGGGAGCGGATTCAGCGCGCTTTTAAAAA 1140
1081 TCTCAGACCGCACTTTGGGAGCGGATTCAGCGCGCTTTTAAAAA 1140
1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174
1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174

RESULT 2
AR075389 1174 bp DNA linear PAT 30-AUG-2006
LOCUS
DEFINITION
ACCESSION AR075389
VERSION AR075389.1 GI:10002139
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1174)
AUTHORS Lund,H. and Kalum,I.
TITLE Garments with considerable variation in abrasion level
JOURNAL Patent: US 5958082-A 1 28-SEP-1999;
FEATURES Location/Qualifiers
source 1..1174
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1174; DB 6; Length 1174;
Best Local Similarity 100.0%; Pred. No. 3.2e-199;
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAGGAGCAGCCCTCAAGCTGACAGTTTCACCCCGCTCTTTTTCGAGCCCGCAGGA 60
1 GAGGAGCAGCCCTCAAGCTGACAGTTTCACCCCGCTCTTTTTCGAGCCCGCAGGA 60
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGAGCTCCG 120
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGAGCTCCG 120
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATCTGAGCACTGGCAAGCCGTCGCTT 180
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATCTGAGCACTGGCAAGCCGTCGCTT 180
181 GAGCCGGGAAAGCCGCGCTCAGCCACCGGCTCAAGCTGCAATGCACTTCCAGGCC 240
181 GAGCCGGGAAAGCCGCGCTCAGCCACCGGCTCAAGCTGCAATGCACTTCCAGGCC 240
241 TGTCCGACTTCATATGCCAGTGGCGGCTGGCAAGCGAGCTGGCCGCTCAGCCAGC 300
241 TGTCCGACTTCATATGCCAGTGGCGGCTGGCAAGCGAGCTGGCCGCTCAGCCAGC 300
301 AGACTCCCTGGCGGCTGAGACGACATCTGCCTCAAGCGCTTCCGCGGAGAGATCCG 360
301 AGACTCCCTGGCGGCTGAGACGACATCTGCCTCAAGCGCTTCCGCGGAGAGATCCG 360
361 GCGGCTCCGAAATCTCTGAGTGGCTGGCGCTGCAAGCCGCTCACTTCGAGTCCG 420
361 GCGGCTCCGAAATCTCTGAGTGGCTGGCGCTGCAAGCCGCTCACTTCGAGTCCG 420

QY 421 TCGCGCGCAAGCAATGCTGCTGCACTGACGACCTGCGGCGACCTGCGAAGTAC 480
DB 421 TCGCGCGCAAGCAATGCTGCTGCACTGACGACCTGCGGCGACCTGCGAAGTAC 480
QY 481 AGTTGATATGCGCATGCGCGCGCGGCGGCGCATCTTCAACGCGCTGAGCTCCAGCT 540
DB 481 AGTTGATATGCGCATGCGCGCGCGGCGGCGCATCTTCAACGCGCTGAGCTCCAGCT 540
QY 541 TCGCGCGCTCCCGCGCGCTCAATACGCGCGCATTTGTGCGCGCAAGCGATTCCT 600
DB 541 TCGCGCGCTCCCGCGCGCTCAATACGCGCGCATTTGTGCGCGCAAGCGATTCCT 600
QY 601 TCGCGCGCGCGCTCAAGCGCGCGCTGCGAGTGGGCTTGAAGCTTCAAGAGCGCGACA 660
DB 601 TCGCGCGCGCGCTCAAGCGCGCGCTGCGAGTGGGCTTGAAGCTTCAAGAGCGCGACA 660
QY 661 ACCGACGTTACGCTTCAAGAGTGGGCTGCGCGCGCGAGATGCTTCCGCGCTCCGCT 720
DB 661 ACCGACGTTACGCTTCAAGAGTGGGCTGCGCGCGCGAGATGCTTCCGCGCTCCGCT 720
QY 721 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 780
DB 721 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 780
QY 781 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 840
DB 781 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 840
QY 841 GTGGCTGCACTGCTCAAGAGTGGGCTGCGAGTGGGCTTCAAGCGGTCGCGT 900
DB 841 GTGGCTGCACTGCTCAAGAGTGGGCTGCGAGTGGGCTTCAAGCGGTCGCGT 900
QY 901 CCGTGTCTCTGCGCACTGCGAGAGTGGGCTTCAAGCGGTCGCGT 960
DB 901 CCGTGTCTCTGCGCACTGCGAGAGTGGGCTTCAAGCGGTCGCGT 960
QY 961 CAGCTTTCGCAAGAGTGGGCTGCGAGAGTGGGCTTCAAGCGGTCGCGT 1020
DB 961 CAGCTTTCGCAAGAGTGGGCTGCGAGAGTGGGCTTCAAGCGGTCGCGT 1020
QY 1021 TTTTGAAGCGCTCAATACATACATACATACATACATACATACATACATACAT 1080
DB 1021 TTTTGAAGCGCTCAATACATACATACATACATACATACATACATACATACAT 1080
QY 1081 TCTCAACCGAGCTTGGGCGCGAGTCAAGCGCGCTTTTAAAAA 1140
DB 1081 TCTCAACCGAGCTTGGGCGCGAGTCAAGCGCGCTTTTAAAAA 1140
QY 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174
DB 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174

RESULT 3
AR094310 1174 bp DNA linear PAT 08-SEP-2000
LOCUS
DEFINITION Sequence 11 from patent US 6001639.
ACCESSION AR094310
VERSION AR094310.1 GI:10021119
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1174)
AUTHORS
Schulein, M., Andersen, L., Nombae, J., Laessen, S., Salsbedren, F.,
Kauppinen, M., Sakari, L., Lange, L., Nielsen, R., Iium, I., Ihara, M., and
Takagi, S.
TITLE
JOURNAL Patent: US 6001639-A, 11-14-DEC-1999;
FEATURES
1. 1174
Location/Qualifiers
source
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1174; DB 6; Length 1174;
Best Local Similarity 100.0%; Pred. No. 3,2e-19;
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGACCCCTCAAGCTGACGCTTCAACCCCGCTCTCTTTCTTGCGGCCAGAGA 60
DB 1 GAGCAGACCCCTCAAGCTGACGCTTCAACCCCGCTCTCTTTCTTGCGGCCAGAGA 60
QY 61 TCGCGCTACCTCCGCTTTCGCAACAACCTGCGCGCTGCACTTCTCTGTCGCTCG 120
DB 61 TCGCGCTACCTCCGCTTTCGCAACAACCTGCGCGCTGCACTTCTCTGTCGCTCG 120
QY 121 CGGCGAGTGCAGTGCAGTCAAGAGATCTGCGAGCTGCGAGCCGTCGTCGCT 180
DB 121 CGGCGAGTGCAGTGCAGTCAAGAGATCTGCGAGCTGCGAGCCGTCGTCGCT 180
QY 181 GCGCGGAGAGGCGCGCTGACCCCAACCGCTTCAAGCGCTGCGAGTCCAGAGCC 240
DB 181 GCGCGGAGAGGCGCGCTGACCCCAACCGCTTCAAGCGCTGCGAGTCCAGAGCC 240
QY 241 TGTCCGACTCAATGTCAGTTCGCGCTGCAAGCGCGCTCGGCTTACTCTGCGCGAC 300
DB 241 TGTCCGACTCAATGTCAGTTCGCGCTGCAAGCGCGCTCGGCTTACTCTGCGCGAC 300
QY 301 AGACTCCCTGCGCGCTGCAAGCAATCTCCCTTCAAGCGCTTCAAGCGAGTTCGCG 360
DB 301 AGACTCCCTGCGCGCTGCAAGCAATCTCCCTTCAAGCGCTTCAAGCGAGTTCGCG 360
QY 361 GCGCGCTGCAATCTGCGCTGCGCGCTGCGAGTTCGCGCTTCAAGCGAGTTCGCG 420
DB 361 GCGCGCTGCAATCTGCGCTGCGCGCTGCGAGTTCGCGCTTCAAGCGAGTTCGCG 420
QY 421 TCGCGCGAGAGCAATGCTGCTGCAAGAGCACTGCGCGAGTTCGCGAGTTCGCG 480
DB 421 TCGCGCGAGAGCAATGCTGCTGCAAGAGCACTGCGCGAGTTCGCGAGTTCGCG 480
QY 481 AGTTGATATGCGCATGCGCGCGCGGCTGCGAGTGGGCTTCAAGCGGTCGCGAGT 540
DB 481 AGTTGATATGCGCATGCGCGCGCGGCTGCGAGTGGGCTTCAAGCGGTCGCGAGT 540
QY 541 TCGCGCGCTCCCGCGCGCTCAATACGCGCGCATTTGTGCGCGCAAGTTCGCT 600
DB 541 TCGCGCGCTCCCGCGCGCTCAATACGCGCGCATTTGTGCGCGCAAGTTCGCT 600
QY 601 TCGCGCGCGCGCTCAAGCGCGCGCTGCGAGTGGGCTTCAAGCGGTCGCGAG 660
DB 601 TCGCGCGCGCGCTCAAGCGCGCGCTGCGAGTGGGCTTCAAGCGGTCGCGAG 660
QY 661 ACCGACGTTACGCTTCAAGAGTGGGCTGCGCGCGCGAGATGCTTCCGCGCTCCGCT 720
DB 661 ACCGACGTTACGCTTCAAGAGTGGGCTGCGCGCGCGAGATGCTTCCGCGCTCCGCT 720
QY 721 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 780
DB 721 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 780
QY 781 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 840
DB 781 GCAAGCGCAAGCACTGCACTTCCCGCTTCAACCGCGCGCAAGCGGTCGCGT 840
QY 841 GTGGCTGCACTGCTCAAGAGTGGGCTGCGAGTGGGCTTCAAGCGGTCGCGT 900
DB 841 GTGGCTGCACTGCTCAAGAGTGGGCTGCGAGTGGGCTTCAAGCGGTCGCGT 900
QY 901 CCGTGTCTCTGCGCACTGCGAGAGTGGGCTTCAAGCGGTCGCGT 960
DB 901 CCGTGTCTCTGCGCACTGCGAGAGTGGGCTTCAAGCGGTCGCGT 960
QY 961 CAGCTTTCGCAAGAGTGGGCTGCGAGAGTGGGCTTCAAGCGGTCGCGT 1020
DB 961 CAGCTTTCGCAAGAGTGGGCTGCGAGAGTGGGCTTCAAGCGGTCGCGT 1020

D6		779	CGAGCCGCCAGTCTCAGGCTCAAGACTCCACGGGCGGTGCATGCTGTGAAGAAGTGCGCTCAAT	838
OY		871	GCGGTGGCATGGCTTTAGGGGATGCAACCACTGTGTCTTCGGACAACACTGGCACAAAT	930
D6		839	GCGGCGGCAATGCTGGAGGGGCTGCAACCACTGGGTGCTGGCAGACACTTGCACAGAGA	898
OY		911	TGAAGCACTACTGCACTGCGAGTGCCTCAAC	961
D6		899	TTTAATACCTGGTACATCACTACATGCTCTGTAGAC	929
RESULT 7				
A23635				
LOCUS				
DEFINITION	A23635	1060 bp	RNA	linear
ACCSSION	A23635			PAT 23-FEB-1995
VERSION	A23635.1			
KEYWORDS				
SOURCE				
ORGANISM	Humicola insolens Humicola insolens Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.			
REFERENCE	1 (bases 1 to 1060)			
ATTORNS	Conventus,A.C., Busch,A. and Baek,A.C.			
TITLE	Defergent compositions with high activity cellulase and softening clays			
JOURNAL	Patent: EP 0495258-A 1 22-JUL-1992;			
THE PROCTER & GAMBLE COMPANY				
FEATURES	Location:Qualifiers			
SOURCE	1..1060			
CDS	/organism="Humicola insolens"			
	/mol_type="unassigned RNA"			
	/db_xref="taxon:3443"			
	10..927			
	/codon_start=1			
	/product="endoglucanase"			
	/protein_id="CAA01681.1"			
	/db_xref="GI:832867"			
	/db_xref="REMITREMBL:CAA01681"			
	/translation="MRSSPLPSAVVAALPYLAADGRSTRVNDCKSGCMAKKAPVQPVFSCNANFORITDPDAKSGCEBGVAAYSCADTPPAVVDDPFAGFPAISINGSNBGMCCACYEELFTSGPPVAGKKWVGSTSTGSDGSNHDLNIPGSGVIGPGCTPOPEGLTGAGGAGTSISRNECDRPDALXEGCYMEFLPMKNANPSEFRVOCCPELVARTGGRINDGNRPVQVPSSSTSPVNPPTSTSTSTSTSPVQPTTPPSGCTAHARMQCGSMGSGCTTVASSTCTKINDIMHQC"			
	10..72			
sig_peptide				
Origin				
Query Match	37.3%; Score 437.4; DB 6; Length 1060;			
Best local Similarity	70.4%; Pred. No. 5.8e-68;			
Matches 655; Conservative	0; Mismatches 246; Indels 30; Gaps 4;			
OY		55	CGAAGATGGCGCTTACTCCGCTTCTTDCACAAACCTCGACGCTGACACTTCTCTGAGTGG	114
D6		5	CAAAGATGGCTTCCCCCTCTCCCGTCCCGGTTGGGCGGCTTGGCGGCTGGCGGATGG	64
OY		115	CTTCGCGGGCCAGGAGGAGTGGCAGTGTCAAGAGATACTGGGGAATGCTGGGAAGCGTGGT	174
D6		65	CCCTTGGCGCTGATGGCAG-----GTCCACCGCTACTGGGAATGTGTGCAAGCTTGTGT	118
OY		175	GCGCTTGGCGCGGANAAGCGCGCGTCAACCGATCTGACGCTGACGATGCAATGCAACTTCC	234
D6		119	GCGGCTGGGCGCAAGAGGCTCCCGTGAACAGACTGTCTTTCTTGCAACGCGCAACTTCC	178
OY		235	AGGGCTGTTCGACCTTCAATGTCTCAGTGGCGTGA---CGGCGGCTCGGCTACATCCCT	291
D6		179	AGGCTATACGGAATCTTGAAGCGCAAGTCCGATCTGCGAGCGGAGGTGTGGCTAACCTGT	238
OY		292	GCGCGCACCAACATCCCTCGGCGGTGAACACACATATCGCTTACAGGCTTGGCGCGACGA	351
D6		239	GCGCGCACCAACAGCCCAATGGGCGTGAACACACACTTGGGCGCTGGGTTTGGTGCACCT	298
OY		352	GCATCGCGCGGCGTCCGAATCTCTGTGTGTGTGTGGCTGTGTACGCGCTTACCTTCACTT	411

Db	239	CTATTGCGGAGCAATGAGGGGGCTGGGTGTCGACCTCTCAACAGAGCTCACTTCAACT	358
QY	412	CGCGTCCCGTGGCGGCAAGACATGGTGGTGTGATCAAGAGCATGGGGGGGACCTGG	471
Db	359	CGCGTCCCGTGGCGGCAAGAGTGGTGGTCACTCAACAGCACTGGCGGTGATCTTG	418
QY	472	GAGTACACAGTTTGATATTCGCAATGCGCGGGGGGCGTGGGCACTTTCCAAAGCGTGGCA	531
Db	419	GCGAGAACCACTTGATCTCAAGATCCCGCGGGGGGGGTGGGCACTTTCAGAGGATGCA	478
QY	532	GCTTCGAGTTGGGGGCTCTCCCGGCGCTCAATACGGGGGCAATTTGTGGCGGACAGT	591
Db	479	CTCCCCAGTTGGGGGCTCTGCCCGGCGAGGCTTACGGGGGCACTGTGTCGGAAAGAGT	538
QY	592	GCGATCTCTTCCCGCGCGCGCTCAAGCGCGGGCTGGCAGTGGGGTGTACATGTTCCAGA	651
Db	539	GCGATGGTTCCCGCGAGCGCTCAAGCGCGGGCTGTACTGGGCTTCCAGCTGGTTGAGA	598
QY	652	ACGCGGACAAACCCGAGGTTTCAGGTTTCAGAGAGTGCAGTCCCGCGAGATGTTGGCC	711
Db	599	ACGCGGACAAATCCGAGGTTTCACTTCCGTGAGGTCAAGTCCGAGCGAGCTGGTGCCTC	658
QY	712	GCTTCGGGCTGGAGCGCAACAGACATCCGAGGTTCCCGGCTTCAACCCCGCAAGGGGAG	771
Db	659	GCAACGGAATGCGCGCGAACAACAAGGCACTTCCGCGGTCAGATCCCTCCAGCA	718
QY	772	GCAACGG-----TGGCACCGGAGACGCGCACGTCGATCGCGCTGGGT	813
Db	719	GCAACAGCTCTCGGTCACACAGGCTTACAGGACACAGACACAGATCCACCAACACT	778
QY	814	CGGCGCAACG---GTCCTCCCGCGCGCGGCAAGTGGTGTGACGTCTCAAGAGTGGGCTCACT	870
Db	779	CGAGCCCGCGAGTCCAGGCTTACAGACTCCGAGCGGCTGTGACGTGAGAGTGGGCTCACT	838
QY	871	GCGGTGGCATGGGCTTCAGCGGATGCAGCACTGGTGTCTGGGCAACCACTCCAGAGAT	930
Db	839	GCGGCGGCAATGGTGGAGCGGCTGCACACACTGGGTGGCGGCAAGCACTTGACGAGACA	898
QY	931	TGAACGACTACTCTGGCAGTGCTCTTAAC	961
Db	899	TTAATGACTGGTACATCACTGATGCCCTGTGAC	929
RESULT 8			
A23644		1060 bp	RNA
LOCUS			linear
DEFINITION	H. insolens endoglucanase gene.		PAT 23-FEB-1995
ACCESSION	A23644		
VERSION	A23644.1		GI:832890
KEYWORDS			
SOURCE	Humicola insolens		
ORGANISM	Humicola insolens		
REFERENCE	Enkayocsa; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.		
AUTHORS	1 (bases 1 to 1060)		
TITLE	Convento,A.C., Busch,A. and Baeck,A.C.		
JOURNAL	Detergent compositions with high activity cellulase and quaternary ammonium compounds		
FEATURES	Patent: EP 0495554-A 1 22-JUL-1992;		
source	THE PROCTER & GAMBLE COMPANY		
	Location/Qualifiers		
	1..1060		
	/organism="Humicola insolens"		
	/mol type="unassigned RNA"		
	/strain="DSM 1800"		
	/db_xref="GI:832891"		
	10..927		
	/codon_start=1		
	/product="endoglucanase"		
	/protein_id="CAA01683.1"		
	/db_xref="GI:832891"		
	/db_xref="REFSEQ:CAA01683"		
	/translation="MRSSPLPAPVAVLPTALADGSTRYWDCKSGCAKAKA		
	VNGVPSCHNANPRTLPDKSGCEBGSVANSQADTPPAVVDPAFGAFTASISGNS		

seq_id	seq	sig_peptide	origin
1	EAAGCQACEELFTSGPVAKKKMWQSTSTGDLGSHNEPDNIIGGGVGFEDGCTPOE GGLEGRDQYVGGISRRBCDFPDALRPGCYKRFDFWRKNAIDPSBPFGVQCPALMAT GGRLNNQDPAVQI.PSSSTSPVDPQSTISTSTSTSTSTSTSTSTSTSTSTSTSTST CGGMSGGCTTCVAGSTCTKINDWYHQC"	10. .72	ORIGIN

Query Match	37.3%;	Score 437.4;	DB 6;	Length 1060;
Best Local Similarity	70.4%;	Pred. No. 5.8e-68;		
Matches 655; Conservative	0;	Mismatches 246;	Indels 30;	Gaps 4;

QY 55 CAGAAATGCGCTCAACTCCGGTCTTTCGCAAAACCTGGGCGCTGCACTTCTCTGGTCTG 114
 Db 5 CCAAGATGTGTTCTCCCTCCCTCTCTCCCTGCTCCGCTGTGTGTGGCGGCTGCGGGGTGG 64
 QY 115 CTTCCGCGGCGAGTGGCGATGGCTACGACAGATATCTGGACTGTGCGAAGCCCTGGT 174
 Db 65 CCGCTTCCGCTGATATGGACG-----GTCCACCCGCTACTAGGAACTGTGCTCAAGCTTCTGT 118
 QY 125 GCGGTGGGCGCGGGAAAGGCGCGCGTCAAGCCAAACCGGCTCTACAGTGTGGATATGCAAACTTCC 234
 Db 119 GCGGCTGGGCGCAAGAAAGGCTCCCGTGAACAGGCTGTCTTTTCTTGCAACGCGAACTTCC 178
 QY 235 AGCGGCTGTCCGACTTTCATATGTCCAGTCCGAGCTGCA---CGGCGGCTCGGCTTACTTCT 291
 Db 179 AGGCTATACGCGAAGCTTTCAGCGCAAGCTCCGGCTCGAGCGGGGCGGTGCGCTTACTCGT 238
 QY 232 GCGCCGCAACGACACTCTCTGGGCGGTGAAACCAACATCTGGCCTTACGCTTCCGCCGCAACA 351
 Db 239 GCGCGCAACGACCCCAATGGGCTGTGAAACCAACATCTGGCGCTGTGTGTGTGTGCGACCT 298
 QY 352 GCATCCGCGGCGGGGTCGGAATCCTGTGTGTGTGTGCGCGCTGACGAGCTCAACCTTCACTT 411
 Db 239 CTATTTGCGCGGCGCAATATGAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 358
 QY 412 CCGGTCCTCGTCCGCGCAAGCAATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471
 Db 359 CCGGTCCTGT 418
 QY 412 GAAATTAACCAAGTATGGAATATGCAATGCAAGCCCGCGCGGCGGTGGATATTTCAACGGCTGCA 531
 Db 419 GCAACCAACCACTTGATCTCAACATATCCCGCGCGCGGTGGATATTTGTGACGAATGCA 478
 QY 532 GCTTCGAGTTTCGAGCGGCTCCCGCGCGCTCAATATCCGAGATTTGTGTGCGCGCAAGT 591
 Db 479 CTCCCGCAAGTTTGGCGGCTGT 538
 QY 552 GCGATTCCTTCCCGCGCGCTCAAGCCCGGCTGCGAGTGGCGGTTGTATGTGTCCAGA 651
 Db 539 GCGATTCGTTTCCCGGAGGCGCTCAAGCCCGGCTGCTACTGCGGCTGTGTGTGTGTGTGT 598
 QY 652 ACGGCGCAACCCGAGGTTTCAAGTTCAGGAGGTTGAGTGGCCCGCGAGATCGTTGACC 711
 Db 559 ACGGCGCAACATCCAGGTTTCAAGTTCGATGAGTGGTCAAGTGGCCAGGAGCTGTGTGTGT 658
 QY 712 GCTTCGCGGCTGCAAGCGCAAGCAAGTCTCAAGTCTTCCGCTTTTCAACCCCGCGAGCGGTT 771
 Db 659 GCAACCGAATGGCGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 718
 QY 712 GCAACG-----TGGCACCGGAGCGCCCAAGTGTGATGTGGCTGTGGGT 813
 Db 719 GCAACGAGCTTCCGATCAACAGACCTTACAGCAACAGCAACCAACCTTCAACCAAGCT 778
 QY 814 CGAGCGCAAGC---GTCTCCCGGCGGCGGAGTGGCTGCAAGCTTCAAGAAATGGGCTAGT 870
 Db 779 CGAGCGCGCAAGTCCAGCTTCAAGCTTCCAGCGGCTGCAAGCTGTGAGAGTGGGCTAGT 838
 QY 871 GCGGTGGCAATCGGCTTCAAGCGGATGCAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 930
 Db 839 GCGGCGGCAATGGCTGTGAGGCGGCTGCAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 898
 QY 931 TGAACGACTTACTGCGAGTGTGCTTAAAC 961

Db	899	TTAATGACTGGTACCATTCAGTGGCTGGAC	929
RESULT 9			
A23953			
LOCUS	A23953	1060 bp	DNA
DEFINITION	Endoglucanase gene.		
ACCESSION	A23953		linear
VERSION	A23953.1	GI:832894	PAT 01-MAR-1995
KEYWORDS			
SOURCE	Humicola insolens		
ORGANISM	Humicola insolens		
REFERENCE	Humicola insolens		
AUTHORS	Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.		
TITLE	1 (bases 1 to 1060)		
JOURNAL	Baack,A.C., Busch,A. and Ceulemans,R.A.A.		
FEATURES	Compact detergent compositions with high activity cellulase		
	Patent: BP 0495257-A 1 22-JUL-1992;		
	THE PROCTER & GAMBLE COMPANY		
	Location/Qualifiers		

CS

```

/organism="Humicola insolens"
/mol_type="unassigned DNA"
/db_xref="taxon:34413"
10..927
/codon_start=1
/product="endoglucanase"
/protein_id="CA01696.1"
/db_xref="GI:832895"
/db_xref="REFRESBL:CA01696"
/translation="MRSSPLPSAVVAALPYLALLADGSRTRYWDCKKSCGAKKAP
VNVQVSCNANFRTLPDPAKSGCEPGVAVASCAADTQPAVNDPFAIGPAATIS
EAGKCAKCELFPTFGSPVAGKXWVQVSTSTGGLGSHNDPAI:EGGVSGLPFGCT
GALPGRYGAGISSNNECDPDPALADKQCYRFPWFLNANLSPSFRVQCPALVAR
GCRANDGSPPAVQIPSSSTSSPVMDPTSTSTSTSTSSPPVQPTTPPSCALENNKA
CGNGSGGCTVAGASTCTXINDWYHQL*"

```

ORIGIN

Query Match	37.3%;	Score 437.4;	DB 6;	Length 1060;
Best Local Similarity	70.4%;	Pred. No. 5.8e-68;		
Matches 655; Conservative	0;	Mismatches 246;	Indels 30;	Gaps 4;

Qy	5	CGAGGATGAGGTCACTCCGGTTTGTGCAACACCTGGACGCTGACATTCTCTGGTGG	114
Db	5	CGAAGTATGGTCTCTCCCTCTCTCCCGTCGGCGTTGTGGCCGCGCTTGGCGGTTGG	64
Qy	115	CCCTCCGCGGACATGGCCAGTGGCCAGTCCACAGATATCTGGAATCTGCGAAGCCGCT	174
Db	65	CCCTTGGCGGTGATGGGAG-----GTCCACCGCGTACTGGGACTGGCGAAGCTTTCGT	118
Qy	175	GCGCTTGGCCCGGGAAAGGCGCGCTGACGCGCAACCGCTCTAAGCGTGGGATGGCACTTCC	234
Db	119	GCGCGTGGGCGCAAGAGGCTCCCGTGAACCAAGCTGTCTTTTCTGCAAGCGCACTTCC	178
Qy	235	AGGCGCTGTCCGACTTCATGTCCAGTCCGGGCTGGCA---CGCGGCTCGGCGCTACTCT	291
Db	179	AGGCTATCAGGACACTTGACACCGAAGCTCGGCGCTGGAGCGCGCGGTTGGCTGACTCGT	238
Qy	292	GCGCGGACCGACATCTCCCTGGCGGGTGAAGACAACTCTGCGCTTACGGCTTGGCGCGACGA	351
Db	239	GCGCGGACCGACACCCCATGGGCTGTGAAGACAACTTGGCGCTTGGCTTGGCGACCT	298
Qy	352	GCAATGCGGCGGAGTCCGAACTCTCGAGGTGCGCGCTCTACGCGGCTACCTCACTT	411
Db	289	CTATTGCGGAGCAATGAAGCGGCGTGGTGTGCGCTCTGTACAGCTCACTTACACT	358
Qy	412	CGGTTCCGTCGCGCGGACAAATGGTGGTGCAGTCAAGACACTGGCGGCGACTTGG	471
Db	359	CGGTTCTGTGTGCTGGGACAAAGATGGTGGTCACTCCACACACTGGCGGATCTTGG	418
Qy	472	GAAGTAAACCAATTCGATTCGCAATGCGCGCGCGCGCGGCGGCAATCTTCAAGGCGTGA	531
Db	419	GCGACACCACTTGATCTCAACTCCCGCGGCGGCGGCGGCGGCAATCTTCAAGGATGGA	478

REFERENCE 1 (bases 1 to 1060)
 AUTHORS Harborts, I.M. and Jansen, M.P.
 TITLE Hydrophobic amines for cellulase stabilization in liquid detergent compositions containing anionic surfactant and cellulase
 JOURNAL Patent: EP 063311-A 1 11-JAN-1995;
 PROCTER & GAMBLE (US)
 COMMENT Other publication CA 2165771 950105.
 FEATURES Location/Qualifiers
 source 1..1060
 /organism="unidentified"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32644"

ORIGIN

Query Match 37.3%; Score 437.4; DB 6; Length 1060;
 Best Local Similarity 70.4%; Pred. No. 5,8e-68;
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

```

QY 55 CCAAGATGCTCTACTCCCTGTTTGGACAAACCTGGCCGCTGACTTCTCTGTGG 114
DB 5 CCAAGATGCTCTACTCCCTGTTTGGACAAACCTGGCCGCTGACTTCTCTGTGG 64
QY 115 CCTCCGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 174
DB 65 CCTTCGCGTGGATGGAG-----GTCCACCGCTACTGGAGTGGAGTGGAGTGG 118
QY 175 GCGCTTGGCGCGGAGAGCGCGCGTGGAGCAACGCTGTAACGCTGGAGTGGCA 234
DB 119 GCGGCTGCGCGAGAGAGCGCTCCGCTGTAACGAGCTGTCTTTCTGCAACCGCA 178
QY 235 AGCGGCTGTGCGACTTAATGTTCAGTGGCGCTGCA-----CGCGCGCTCGG 291
DB 179 AGCGATCAAGGACTTGAAGCGAGTGGCGCTGGAGCGCGCGTGTGCTGCTACT 238
QY 292 GCGCGGAGCAAGCTCCCTGGCGGAGAGCAACATCTCGCTACGCGCTTGGCGGAG 351
DB 239 GCGCGGAGCAAGCTCCCTGGCGGAGAGCAACATCTCGCTACGCGCTTGGCGGAG 298
QY 352 GCAATGCGCGGCGGAGTCCGAATCTCGTGGTGGTGGTGGTGGTGGTGGTGGT 411
DB 299 CTATTGCGCGGAGCAATGAGAGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 358
QY 412 CGGTCGCGTGGCGGAGAGCAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 471
DB 359 CGGTCGCTGTGGAGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 418
QY 472 GAAGTACCAAGTTCGATATGCGCATGCGCGCGCGCGCGCGCATCTTCAACGCG 531
DB 419 GAGGCAACCACTGATCTCAATCCCGGCGCGCGCGCGCGCATCTTCAACGAGTGA 478
QY 532 GCTGCAAGTGGCGGCTCCCGCGCGCTCAATGCGCGCGCATTTGCTGGCGAGCA 591
DB 479 CTCGCCAGTTGGCGGCTGTGCGCGCGCGCGCGCATCTTGTCCGGAAGAGT 538
QY 592 GCGATTCCTTCCCGCGCGCTCAAGCGCGCGCTGCGAGTGGCGGTTTGAAGTTC 651
DB 539 GCGATGCGTTCGCGAGCGCTCAAGCGCGCGCTGCTAAGCGCGCTTGAAGTTC 598
QY 652 AGCGCGCAACCGAGCTTCAAGCTTCAAGAGTGGAGTGGAGTGGAGTGGAGTGG 711
DB 599 AGCGCGCAACCGAGCTTCAAGCTTCAAGAGTGGAGTGGAGTGGAGTGGAGTGG 658
QY 712 GCTTCGCGCTGAGAGCGAGAGTCAAGTCAAGTCCCGCTTTCACCCCGCAAGG 771
DB 659 GCAACCGAGTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
QY 772 GCAAGG-----TGGACGCGGAGAGCGCGCAAGTTCAGCTGGCGTGGGT 813
DB 719 GCAACAGCTTCGCGTCAACAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 778
QY 814 GGGGCGAGAC---GTCTCCCGCGCGCGAGTGGCTGCAAGTCTGAGAGTGGGCT 870
DB 779 CGAGCCGCGAGTCCAGCTCAAGACTCCAGCGGCTGCACTGTGAGAGTGGGCT 838
  
```

QY 871 GCGGTGGCATGCGCTTACGCGGAGTACCACTGTGTCTGTGGACCACTGGCAGAGT 930
 DB 839 GCGCGGCAATGCGTGGAGCGCGGAGCAACCTGGCTGGTGGAGCACTTGGACAGAA 898

QY 931 TGAACGACTACTGCGAGTGGCTCTTAAC 961
 DB 839 TTAATGACTGTTCATCAAGTGGCTGTAGAC 929

RESULT 12
 A68060 1060 bp DNA linear PAT 05-MAY-1999
 LOCUS A68060
 DEFINITION Sequence 55 from Patent W09743409.
 ACCESSION A68060
 VERSION A68060.1 GI:4756862
 KEYWORDS
 SOURCE Humicola insolens
 ORGANISM Humicola insolens
 Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.
 REFERENCE 1 (bases 1 to 1060)
 AUTHORS Dalboege, H., Diderichsen, B., Sandal, T. and Kauppinen, S.
 TITLE METHOD OF PROVIDING NOVEL DNA SEQUENCES
 JOURNAL Patent: WO 9743409-A 55 20-NOV-1997;
 NOVONORDISK AS (DK)
 FEATURES Location/Qualifiers
 source 1..1060
 /organism="Humicola insolens"
 /mol_type="unassigned DNA"
 /strain="DSM 1800"
 /db_xref="taxon:34413"
 10..927
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CA842307.1"
 /db_xref="GI:4756863"
 /translation="MRSSPLPSAVVAALPVALAANDGRSTRVMOCKPSGMAKKAP
 VNPVPSGNANFORITDPDKSGCTERGYNASCDQTPHAYNDPLSPATISASGN
 EAGMCCACVETLPTSGPVAKRWVOSTSGDLSNHEPLNTPGGGVTGDECTP
 GAGPQRYGELISNRECDRPPDLAKGCMFDFWPNXADPFSFROVOCDAELVAT
 GGRNDGNFPAVOIPSSSTSPVQPTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST
 CGNGSGGTTCTGAGTCTKINDWHQCL"
 10..72
 sig_peptide
 mat_peptide
 73..924
 /product="unnamed"

ORIGIN

Query Match 37.3%; Score 437.4; DB 6; Length 1060;
 Best Local Similarity 70.4%; Pred. No. 5,8e-68;
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

```

QY 55 CCAAGATGCTCTACTCCCTGTTTGGACAAACCTGGCCGCTGACTTCTCTGTGG 114
DB 5 CCAAGATGCTCTACTCCCTGTTTGGACAAACCTGGCCGCTGACTTCTCTGTGG 64
QY 115 CCTCCGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 174
DB 65 CCTTCGCGTGGATGGAG-----GTCCACCGCTACTGGAGTGGAGTGGAGTGG 118
QY 175 GCGCTTGGCGCGGAGAGCGCGTGGAGCAACGCTGTAACGCTGGAGTGGCA 234
DB 119 GCGGCTGCGCGAGAGAGCGCTCCGCTGTAACGAGCTGTCTTTCTGCAACCGCA 178
QY 235 AGCGGCTGTGCGACTTAATGTTCAGTGGCGCTGCA-----CGCGCGCTCGG 291
DB 179 AGCGATCAAGGACTTGAAGCGAGTGGCGCTGGAGCGCGCGTGTGCTGCTACT 238
QY 292 GCGCGGAGCAAGCTCCCTGGCGGAGAGCAACATCTCGCTACGCGCTTGGCGGAG 351
DB 239 GCGCGGAGCAAGCTCCCTGGCGGAGAGCAACATCTCGCTACGCGCTTGGCGGAG 298
QY 352 GCAATGCGCGGCGGAGTCCGAATCTCGTGGTGGTGGTGGTGGTGGTGGTGGT 411
  
```


Patkar, S. Anant., Hagen, F., Hjort, C. Mailand, and Hestrup, S.
Cellulase preparation comprising an endoglucanase enzyme
JOURNAL Patent: US 5948672-A 1 07-SEP-1999;
FEATURES Location/Qualifiers
source 1..1060
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 37.3%; Score 437.4; DB 6; Length 1060;

Best Local Similarity 70.4%; Pred. No. 5.8e-68;

Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

55 CCAGATGCGCTCTACTCCCGTTCTTCCGACAAACCTGGCCGCTGCACTTCTGTGCG 114
|||
5 CCAGATGCGCTCTACTCCCGTTCTTCCGACAAACCTGGCCGCTGCACTTCTGTGCG 64
|||
115 CCTCGCGGCGAGTGGGAGTGGCGAGTCCAGAGATATAGGAGATGCGAGACCGTGT 174
|||
65 CCTTGGCGCTGATGGAG-----GTCCACCGGCTACTGGGACTGCTGCAAGCTTGT 118
|||
175 GCGCTTGGCCCGGAGAGCGCGCTGACGACCAACCGGTCTACCGGTGGAATGCCACTTCC 234
|||
119 GCGGCTGGGCGAGAGAGGCTCCCGTGAACGAGCTGTCTTCTTGGAAGGCCAATCTTC 178
|||
235 AGCGCTGTCCGATTTCAATGTCCAGTGGGCTTGCA---CGGCGGCTGGCGCTACTCT 291
|||
179 AGGATACAGAGATTCGAGAGCGCAAGTCGCGCTGCGAGCGCGGCGGTGTGCGCTACTCT 238
|||
292 GCGCGGACAGATCCCTGGGCGGTGAGACGACATCTGCTCAACCGCTTCCGCGGACGA 351
|||
239 GCGCGGACAGATCCCTGGGCGGTGAGAGCGAGCTGTGCGCTGCTGCTGCTGCTGCT 298
|||
352 GCAATGCGCGGCGGTGCGAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 411
|||
299 CTATGCGCGGCGAGAGAGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 358
|||
412 CCGGTCCGCTGCGCGGAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471
|||
359 CCGGTCTGT 418
|||
472 GAAGTACAGATTCGATATGCGCATGCGCGGCGGCGGTGAGCATCTTCAACGAGTGA 531
|||
419 GCGAGAACATCTTGATATCAATCTCCGCGGCGGCGGTGAGCATCTTCAACGAGTGA 478
|||
532 GCTTCAGATTCGCGGCTCTCCGCGGCTCTCAATACGCGGCGCATTTGTGTGTGTGT 591
|||
479 CTCCCAAGTTGCGCGGT 538
|||
592 GCGATTCCTGCGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 651
|||
539 GCGATTCCTGCGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 598
|||
652 AGCGGACAGACCGGAGCTTCAAGTTTCAGAGGTGAGTGTGTGTGTGTGTGTGTGTGT 711
|||
599 AGCGGACAGACCGGAGCTTCAAGTTTCAGAGGTGAGTGTGTGTGTGTGTGTGTGTGT 658
|||
712 GCTTCGCGGTGAGGCGGAGAGCTCAAGTCTCAAGTCTTCAACCGCGGCGAGGCTG 771
|||
659 GCGCGGAGTTCGCGGCGGAGAGAGCGGAGCATTTCTGTGTGTGTGTGTGTGTGTGT 718
|||
772 GCGAGG-----TGGCAGCGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 813
|||
719 GCGAGGCTCTCCGCTCAACGAGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
|||
814 GCGGCGAGAG---GTCTCCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
|||
779 GAGGCGCGGAG 838
|||
871 GCGGCGGAG 930
|||
839 GCGGCGGAG 898
|||

Qy 931 TGACGACTACTCTGCGAGGCGCTGAAAC 961
|||
Db 899 TTATGACTGTGATCATCATGAGCTGTGAC 929
|||

RESULT 15

AR163160

LOCUS AR163160 1060 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 55 from patent US 6270968.

ACCESSION AR163160

VERSION AR163160.1 GI:16233662

KEYWORDS

SOURCE Unknown.

ORGANISM

REFERENCE 1 (bases 1 to 1060)

AUTHORS Dahl, O. Blashed, G. H., Sandal, T., Kauppinen, M. Sakari, and

TITLE Method of providing a hybrid polypeptide exhibiting an activity of

JOURNAL Patent: US 6270968-A 55 07-AUG-2001;

FEATURES

source

ORIGIN

Query Match 37.3%; Score 437.4; DB 6; Length 1060;

Best Local Similarity 70.4%; Pred. No. 5.8e-68;

Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

55 CCAGATGCGCTCTACTCCCGTTCTTCCGACAAACCTGGCCGCTGCACTTCTGTGCG 114
|||
5 CCAGATGCGCTCTACTCCCGTTCTTCCGACAAACCTGGCCGCTGCACTTCTGTGCG 64
|||
115 CCTCGCGGCGAGTGGGAGTGGCGAGTCCAGAGATATAGGAGATGCGAGACCGTGT 174
|||
65 CCTTGGCGCTGATGGAG-----GTCCACCGGCTACTGGGACTGCTGCAAGCTTGT 118
|||
175 GCGCTTGGCCCGGAGAGCGCGCTGACGACCAACCGGTCTACCGGTGGAATGCCACTTCC 234
|||
119 GCGGCTGGGCGAGAGAGGCTCCGATGAACAGCTGTCTTCTGTGCAAGCGCACTTCC 178
|||
235 AGCGGCTGT 291
|||
179 AGCGGCTGT 238
|||
292 GCGCGGACAGATCCCTGGGCGGTGAGACGACATCTGCTCAACCGCTTCCGCGGACGA 351
|||
239 GCGCGGACAGATCCCTGGGCGGTGAGAGCGAGCTGTGCGCTGCTGCTGCTGCTGCT 298
|||
352 GCAATGCGCGGCGGTGCGAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 411
|||
299 CTATGCGCGGCGAGAGAGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 358
|||
412 CCGGTCCGCTGCGCGGAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471
|||
359 CCGGTCTGT 418
|||
472 GAAGTACAGATTCGATATGCGCATGCGCGGCGGCGGTGAGCATCTTCAACGAGTGA 531
|||
419 GCGAGAACATCTTGATATCAATCTCCGCGGCGGCGGTGAGCATCTTCAACGAGTGA 478
|||
532 GCTTCAGATTCGCGGCTCTCCGCGGCTCTCAATACGCGGCGCATTTGTGTGTGTGT 591
|||
479 CTCCCAAGTTGCGCGGT 538
|||
592 GCGATTCCTGCGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 651
|||
539 GCGATTCCTGCGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 598
|||
652 AGCGGACAGACCGGAGCTTCAAGTTTCAGAGGTGAGTGTGTGTGTGTGTGTGTGTGT 711
|||

Db 539 AGGCCGACAAATCCGAGCTTCAAGCTTCGGTCAGCTCAAGTCCAGCCGAGCTCTGCTC 658
Qy 712 GCTCGGCTGCAGAGCGCAACGACTTCAGCTTCCCGCTTTCACCCCCCAAGCGGTG 771
Db 659 GCACCGAATGCGCGGCAACGACGCGCACTTCCCTGCGGTCCAGATCCCTTCAGCA 718
Qy 772 GCMAAGG-----TGGACCGGGAAGCCCAAGTCACTGCGCTGAGT 813
Db 719 GCACCAAGCTTCGCGTCAACCAAGCTTACGACACCAAGCACCAAGTCCACCTCACT 778
Qy 814 CGGACCAAGC---GTCTCCGCGCGGCGAGTGGCTGCAAGCTTCAGAAATGGGCTCAGT 870
Db 779 CGAGCCGCGCAATCCAGCTTACGACTCCCAAGCGGCTGCACTGCTGAAGAGTGGGCTCAGT 838
Qy 871 GCGGTGGCATGCGCTTCAAGCGGATGCAACCACTGTGTCTTGGACACCCTGCAAGAT 930
Db 839 GCGCGGCAATGCGTGAAGCGGCTGCAACCACTGCGTGGTGGAGCACTTGCAGAAAGA 898
Qy 931 TGAAGCACTACTGCTGCGAGTGGCTTAAAC 961
Db 899 TTAATGACTGGTACCATGAGTGCCTGTAGAC 929

Search completed: July 7, 2004, 10:19:30
Job time : 7435 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compygen Ltd.

OM nucleic - nucleic search, using *sw* mode:

Run on: July 7, 2004, 08:03:07 ; Search time 716 Seconds
(without alignments)
6965.621 Million cell updates/sec

Title: US-10-007-521-11

Perfect score: 1174
Sequence: 1 GAGGACGACCCCTCAAGCTG.....AAAAAAAAAAAAAAAAAAAA 1174

Scoring table: IDENTITY NTC
Gapop 10.0, Gapext 1.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Geneseg 29Jan04:*
2: geneseg1980s:*
3: geneseg1990s:*
4: geneseg2000s:*
5: geneseg2001as:*
6: geneseg2001bs:*
7: geneseg2002as:*
8: geneseg2003bs:*
9: geneseg2003cs:*
10: geneseg2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1174	100.0	1174	2	AAT39050
2	1174	100.0	1174	2	AAT39050
3	453.6	38.6	1261	2	AAV39096 Monocompo
4	451.6	38.5	922	2	AAV23748 Humicola
5	439	37.4	1060	2	AAV15072 Hybrid DN
6	437.4	37.3	1058	2	AAQ26405 Humicola
7	437.4	37.3	1060	2	AAQ14856 Humicola
8	437.4	37.3	1060	2	AAQ25932 Cellulase
9	437.4	37.3	1060	2	AAQ30067 Sequence
10	437.4	37.3	1060	2	AAQ29934 Endoglucanase
11	437.4	37.3	1060	2	AAQ49941 Endoglucanase
12	437.4	37.3	1060	2	AAQ41732 Dye trans
13	437.4	37.3	1060	2	AAQ10182 Alkaline
14	437.4	37.3	1060	2	AAV15065 Humicola
15	437.4	37.3	1060	2	AAV16102 Humicola
16	437.4	37.3	1060	2	AAV16102 Humicola
17	435.8	37.1	1060	2	AAQ26380 Endoglucanase
18	431	36.7	1060	2	AAQ30072 43RD endo
19	429.4	36.6	1060	2	AAQ31181 H. insol
20	412.8	35.2	922	2	AAV15073 Hybrid DN
21	393.4	33.5	925	2	AAV15076 Hybrid DN
22	390.6	33.3	807	2	AAV16104 Humicola
23	377.2	32.1	1257	2	AAV19378 Humicola

24	377.2	32.1	1257	2	AAV13840
25	377.2	32.1	1257	6	AAV43252
26	377.2	32.1	1257	6	AAV60700
27	377.2	32.1	1257	6	AAV59235
28	377.2	32.1	1257	8	AAV37115
29	368.2	31.4	1154	2	AAV139048
30	366.4	31.2	928	2	AAV15074
31	365.6	31.1	927	2	AAV29596
32	353.6	30.1	936	2	AAV6541
33	337.6	28.8	915	2	AAV15075
34	337	28.7	894	2	AAV139061
35	334.8	28.5	927	2	AAV139062
36	315.2	26.8	1473	2	AAQ14857
37	315.2	26.8	1473	2	AAQ25933
38	315.2	26.8	1473	2	AAQ26407
39	315.2	26.8	1473	2	AAQ26382
40	315.2	26.8	1473	2	AAQ29935
41	315.2	26.8	1473	2	AAQ49942
42	315.2	26.8	1473	2	AAQ60179
43	315.2	26.8	1473	2	AAV16103
44	313.6	26.7	885	2	AAV139075
45	313.6	26.7	1473	2	AAQ41733

ALIGNMENTS

RESULT 1	
AAV39050	
ID	AAT39050 standard; cDNA; 1174 BP.
XX	
AC	AAT39050:
XX	
DT	20-MAY-1997 (first entry)
XX	
DE	cDNA encoding cellulytic enzyme #4 of the invention.
XX	
KM	Cellulytic enzyme; endoglucanase; hydrolisis; cellulose; microorganism;
KM	plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KM	stone-washing; cellulosic fabric; colour clarification; delibration;
KM	cell wall degradation; paper pulp; debarking; fibre modification;
KM	enzymatic de-inking; drainage improvement; ss.
XX	
OS	Thielavia terrestris.
XX	
FH	Key
FT	Location/Qualifiers
FT	59.959
FT	/*tag= a
FT	/product= "endoglucanase"
XX	
WM	W09629397-A1.
XX	
PD	26-SEP-1996.
XX	
PF	18-MAR-1996; 96MO-DK000105.
XX	
PR	17-MAR-1995; 95DK-00000272.
PR	08-AUG-1995; 95DK-00000885.
PR	08-AUG-1995; 95DK-00000886.
PR	08-AUG-1995; 95DK-00000887.
PR	08-AUG-1995; 95DK-00000888.
PR	12-FEB-1996; 96DK-00000137.
XX	
PA	(NOVO) NOVO-NORDISK AS.
XX	
PI	Schuelein M, Andersen LN, Laessen SF, Kauppinen MS, Lange L;
XX	Nielsen R, Ihara M, Takagi S;
DR	WPI: 1996-443173/44.
XX	P-PSDB; AAM04928.
XX	New endoglucanase enzyme preparations - contg. conserved catalytic
PT	regions, useful for treating fabrics, textiles, plant material or paper

PT pulp.
 XX Claim 72; Page 160-161; 316pp; English.
 XX
 CC AAT3047-139054 represent the coding sequences for the enzymes of the
 CC invention. The encoded enzymes possess cellulolytic (particularly
 CC endoglucanase) activity. Cellulolytic enzymes are involved in the
 CC hydrolysis of cellulose, and are synthesised by a large number of
 CC microorganisms and plants. The enzymes of the invention containing the
 CC conserved catalytic regions (such as AAM04913) exhibit improved
 CC performance, e.g. 50 times higher performance, compared to multiple
 CC domain enzymes. The enzymes can be used for the treatment of fabrics or
 CC textiles, preferably for preventing backstaining, for bio-polishing or
 CC for stone-washing cellulosic fabric. They can also be used to provide
 CC colour clarification for laundry. The enzymes can also be used for the
 CC degradation or modification of plant material, such as cell walls. They
 CC can also be used in the treatment of paper pulp preferably for debarking,
 CC defibration, fibre modification, enzymatic de-inking or drainage
 CC improvement
 CC
 XX
 SQ Sequence 1174 BP; 243 A; 395 C; 320 G; 216 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1174; DB 2; Length 1174;
 Best Local Similarity 100.0%; Pred. No. 9e-204;
 Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGGAGCAACCCCTCAAGCTGTAAGTTTCAACCCGCTCTCTTTCTGCGCCCAAGAA 60
 DB 1 GAGGAGCAACCCCTCAAGCTGTAAGTTTCAACCCGCTCTCTTTCTGCGCCCAAGAA 60
 QY 61 TGGGCTCTACTCCCGCTTCTTGCAACCCGCGCTGCACTTCCCTGCTGCTCCG 120
 DB 61 TGGGCTCTACTCCCGCTTCTTGCAACCCGCGCTGCACTTCCCTGCTGCTCCG 120
 QY 121 CGGCGAGTGGGAGTGGGCAAGTCAAGATCTGATCTGCAAGCCGCTGCTG 180
 DB 121 CGGCGAGTGGGAGTGGGCAAGTCAAGATCTGATCTGCAAGCCGCTGCTG 180
 QY 121 CGGCGAGTGGGAGTGGGCAAGTCAAGATCTGATCTGCAAGCCGCTGCTG 180
 DB 121 CGGCGAGTGGGAGTGGGCAAGTCAAGATCTGATCTGCAAGCCGCTGCTG 180
 QY 181 GGGCGGGAAGGCGCGCTGAGCAACCCGCTGCAAGCCGCTGCAAGCCGCTG 240
 DB 181 GGGCGGGAAGGCGCGCTGAGCAACCCGCTGCAAGCCGCTGCAAGCCGCTG 240
 QY 241 TGTCCGACTTCAATGTCAGTGGCTGCAAGCCGCTGCAAGCCGCTGCAAGCC 300
 DB 241 TGTCCGACTTCAATGTCAGTGGCTGCAAGCCGCTGCAAGCCGCTGCAAGCC 300
 QY 301 AGACTCCCTGGGCGGAGCAATCTGCTGCAAGCCGCTGCAAGCCGCTGCAAGCC 360
 DB 301 AGACTCCCTGGGCGGAGCAATCTGCTGCAAGCCGCTGCAAGCCGCTGCAAGCC 360
 QY 361 GGGGCTCCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 DB 361 GGGGCTCCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 421 TGGCGGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 421 TGGCGGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 481 AGTTGATATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 DB 481 AGTTGATATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 541 TGGGCGGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 541 TGGGCGGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 601 TGGCGGCGGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 DB 601 TGGCGGCGGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 661 ACCGAGGTTCAAGCTTCAAGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 720
 DB 661 ACCGAGGTTCAAGCTTCAAGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 720

QY 721 GCAAGCGCAAGCAAGCAAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 780
 DB 721 GCAAGCGCAAGCAAGCAAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 780
 QY 781 GCAAGCGCAAGCAAGCAAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 840
 DB 781 GCAAGCGCAAGCAAGCAAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 840
 QY 841 GTGGCTGCAAGCTTCAAGAGTGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 900
 DB 841 GTGGCTGCAAGCTTCAAGAGTGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 900
 QY 901 CTTGCTGCTGCAAGCTTCAAGAGTGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 960
 DB 901 CTTGCTGCTGCAAGCTTCAAGAGTGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 960
 QY 961 CACTTTTCGCAAGAGTGGCGGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCA 1020
 DB 961 CACTTTTCGCAAGAGTGGCGGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCA 1020
 QY 1021 TTTTGAAGGCTCAATACATACATACATACATACATACATACATACATACAT 1080
 DB 1021 TTTTGAAGGCTCAATACATACATACATACATACATACATACATACATACAT 1080
 QY 1081 TCTCAACCGCACTTGGGCGGCAAGTCAAGGCGCTTTTAAAAA 1140
 DB 1081 TCTCAACCGCACTTGGGCGGCAAGTCAAGGCGCTTTTAAAAA 1140
 QY 1141 AA 1174
 DB 1141 AA 1174
 RESULT 2
 AA39096
 ID AA39096 standard; cDNA; 1174 BP.
 AC AA39096;
 DT 21-SEP-1998 (first entry)
 XX
 DE Monocomponent endoglucanase encoding cDNA.
 XX
 KM Monocomponent endoglucanase; garment; abrasion;
 KM abraded looking jeans; fungus; Thielavia terrestris; stone-washed; ss.
 XX
 OS Thielavia terrestris.
 XX
 FH Key Location/Qualifiers
 PT CDS 60..959
 FT /tag="a
 FT /product="monocomponent endoglucanase"
 XX
 PN EP843041-A1.
 XX
 PD 20-MAY-1998.
 XX
 PF 30-MAY-1997; 97EP-00610021.
 XX
 PR 13-NOV-1996; 96DK-00001276.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Lund H, Kalum L;
 XX
 DR WPI; 1998-263308/24.
 XX
 PT P-PSDB; AAM63624.
 XX
 PS Pair of jeans - with abraded look.
 XX
 Claim 17; Page 9-10; 16pp; English.

CC This cDNA encodes a monocomponent endoglucanase. This is a cellulolytic
CC enzyme belonging to the fungal species *Thielavia terrestris*. This can be
CC used in the process of invention of providing a pair of jeans made from
CC dyed twill fabric that has localised variations in the colour density of
CC the fabric. The process involves providing the jeans with a stone-washed
CC or abraded look where the variation corresponds to a delta remission
CC value (delta R) higher than 11, and the reflection of a first area of the
CC jeans fabric is less than 12 percent. The process comprises selecting the
CC desired textile cutting pattern for jeans garment, positioning the
CC pattern onto newly manufactured dyed twill fabric, cutting the jeans
CC garment parts, sewing the pair of jeans, optionally subjecting the pair
CC of jeans to a desizing treatment, and subjecting the pair of jeans to an
CC abrasion treatment with an efficient amount of a cellulolytic enzyme in
CC an aqueous medium essentially free of bleaching chemicals. The abraded
CC look near stitching is different from the abraded look far from stitching
XX
SQ Sequence :174 BP; 243 A; 395 C; 320 G; 216 T; 0 U; 0 Other;

Query Match 100.0%; Score 1174; DB 2; Length 1174;
Best Local Similarity 100.0%; Pref. No. 9e-204;
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGACCCCTCAAGCTGTACAGTTCCACCCGCTCTTTCTTCCGCCCCCAGAGA 60
DB 1 GAGCAGACCCCTCAAGCTGTACAGTTCCACCCGCTCTTTCTTCCGCCCCCAGAGA 60
QY 61 TCGGCTTACTCCGCTTCTTGGCAACCTTGGCGCTGCACTTCTTGTGCTCCG 120
DB 61 TCGGCTTACTCCGCTTCTTGGCAACCTTGGCGCTGCACTTCTTGTGCTCCG 120
QY 121 CGGCGAGTGGAGTGGCGAGTCCAGAGATCGGAGACGCGTGAAGCGGTGGGCTT 180
DB 121 CGGCGAGTGGAGTGGCGAGTCCAGAGATCGGAGACGCGTGAAGCGGTGGGCTT 180
QY 181 GGGCCGGGAAAGCCCGCGCTCAGCACAACCGGCTCAGCGATGCGAATCTTCAGAGCC 240
DB 181 GGGCCGGGAAAGCCCGCGCTCAGCACAACCGGCTCAGCGATGCGAATCTTCAGAGCC 240
QY 241 TGTCCGACTTCAATGTCTGAGTGTGCTGCAACCGGCGCTGCGCTTCTTGTGCGAAC 300
DB 241 TGTCCGACTTCAATGTCTGAGTGTGCTGCAACCGGCGCTGCGCTTCTTGTGCGAAC 300
QY 301 AGACTCCCTGGGCGGTGAGACGACATCTCCGCTACGCGTGGCGAGAGATGCGCG 360
DB 301 AGACTCCCTGGGCGGTGAGACGACATCTCCGCTACGCGTGGCGAGAGATGCGCG 360
QY 361 GCGGGTCCGATCTCTGTGTGCTGCGCTGCTACGCGCTCACTTCCGCTCCG 420
DB 361 GCGGGTCCGATCTCTGTGTGCTGCGCTGCTACGCGCTCACTTCCGCTCCG 420
QY 421 TCGCCGCGAAGACATGTGTGTGCTGCTGCAAGCCTGCGCGAGAAATGAC 480
DB 421 TCGCCGCGAAGACATGTGTGTGCTGCTGCAAGCCTGCGCGAGAAATGAC 480
QY 481 AGTTGATATGCGCATGCGCGCGCGGCGGTGGGCACTTTCAACGCGCTGCGAGT 540
DB 481 AGTTGATATGCGCATGCGCGCGCGGCGGTGGGCACTTTCAACGCGCTGCGAGT 540
QY 541 TCGCGGCGCTCCCGCGCTCAATACGCGGCAATTTGTGCGCGAGCAGTGTGCT 600
DB 541 TCGCGGCGCTCCCGCGCTCAATACGCGGCAATTTGTGCGCGAGCAGTGTGCT 600
QY 601 TCGCGGCGCGCTGCAAGCGCGCTGCGCGAGTGTGCTGCTGCTGCTGCTGCT 660
DB 601 TCGCGGCGCGCTGCAAGCGCGCTGCGCGAGTGTGCTGCTGCTGCTGCTGCT 660
QY 661 ACCGAGGTTGAGTTTCAGAGTGTGAGTGTGCGCGAGATGTTGCGCGCT 720
DB 661 ACCGAGGTTGAGTTTCAGAGTGTGAGTGTGCGCGAGATGTTGCGCGCT 720
QY 721 GCAAGCGCAAGAGAGCTCAGCTTCCGCTTTCACCCCGCAAGGCTGTGAGGCT 780
DB 721 GCAAGCGCAAGAGAGCTCAGCTTCCGCTTTCACCCCGCAAGGCTGTGAGGCT 780

QY 781 GCACCGGAGCGCCACGTCGATCGCCCTGGGTGGGCGAGAGTCTCCCGGCGCGCA 840
DB 781 GCACCGGAGCGCCACGTCGATCGCCCTGGGTGGGCGAGAGTCTCCCGGCGCGCA 840
QY 841 GTGGCTGCAAGTGTGAGAGTGGGCTCAGTGTGGGCTATCGGCTTCAGCGGATGACCA 900
DB 841 GTGGCTGCAAGTGTGAGAGTGGGCTCAGTGTGGGCTATCGGCTTCAGCGGATGACCA 900
QY 901 CCGTGTCTTGGCACAACCTGCGAGAGTGAAGCACTACTGCGAGTCCCTTAA 960
DB 901 CCGTGTCTTGGCACAACCTGCGAGAGTGAAGCACTACTGCGAGTCCCTTAA 960
QY 961 CAGTTTGTGCAAGGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020
DB 961 CAGTTTGTGCAAGGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020
QY 1021 TTTTGGAGGCTCAATACATACATTAATCTTGTATGATACAGCGCGGTACACA 1080
DB 1021 TTTTGGAGGCTCAATACATACATTAATCTTGTATGATACAGCGCGGTACACA 1080
QY 1081 TCTCAGACCGACTTGGGCGGAGATCAGGCGGCTTTTAAAAA 1140
DB 1081 TCTCAGACCGACTTGGGCGGAGATCAGGCGGCTTTTAAAAA 1140
QY 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174
DB 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174

RESULT 3
AAV23748
ID AAV23748 standard; cDNA, 1261 BP.
XX
XX AAV23748;
XX
DT 19-AUG-1998 (first entry)
DE Humicola grisea cellulase cDNA sequence.
XX
XX Cellulase; variant; cellulolytic enzyme; endoglucanase V; E5V; paper;
XX protein engineering; detergent; fabric softener; pulp; debarking;
XX defibration; fibre modification; degradation; ss.
OS Humicola grisea.
XX
XX MO9812307-A1.
XX
XX PD 26-MAR-1998.
XX
XX PF 17-SEP-1997; 97MO-DK000393.
XX
XX PR 17-SEP-1996; 96DK-00001013.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
XX PI Andersen KV, Christiansen L, Damgaard B;
XX
XX WPI; 1998-217251/19.
XX
XX DR
XX PT Cellulase enzyme variants - having amino acid changes which improve
XX properties e.g. activity, sensitivity to surfactants, pH optimum or
XX stability.
XX
XX PS Example 3; Page 77-78; 115pp; English.
XX
XX CC The present invention describes a cellulase enzyme variant comprising a
XX catalytic core domain exhibiting cellulolytic activity which is derived
XX from a naturally occurring parental cellulase by amino acid residue
XX substitution, insertion and/or deletion; and (with cellulase numbering):
XX (a) at position 5 has an Ala, Ser, or Thr residue; (b) at position 8 has
XX a Phe or a Tyr residue; (c) at position 9 has a Phe, Trp, or Tyr residue;
XX (d) at position 10 has a Asp residue; (e) at position 121 has a Asp

CC residue. The present sequence represents the cDNA sequence of Humicola
 CC grisea cellulase from the present invention. The new cellulase enzymes
 CC can be used in e.g. detergent or fabric softener compositions, for bio-
 CC polishing of new fabrics, for promoting a stone-washed look to cellulosic
 CC containing fabric, for pulp and paper applications, e.g. for debarking,
 CC defibrillation, fibre modification, drainage improvement, inter fibre
 CC bonding or for degradation of plant material e.g. for improving feed
 CC value. The cellulase variants have improved properties with respect to
 CC e.g. catalytic activity, altered sensitivity to anionic tensides, pH
 CC optimum or activity profile or stability

XX Sequence 1261 BP; 246 A; 417 C; 336 G; 262 T; 0 U; 0 Other;

Query Match 38.6%; Score 453.6; DB 2; Length 1261;

Best Local Similarity 74.9%; Pred. No. 3.8e-73;

Matches 583; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

26 TTTCACCCCGCTCTCTTTCTTGGGCCCCGAGAGAGGCGCTTAACTCCGCTCTTGGAC 85
 18 TTTTATTCAGCTCATTTATTTTAAACTTCAATATGCGCTGCTCTTAATTTTCGGAC 77
 86 AACCTGCGCGCTGCACTTCTCTGCTGCTCCGCGGCAATGAGAGTGGCACTGCAC 145
 78 GGCCTTGGCGGCTGCGCTCTCCCTTGCAGACATCGCGCGGATGGCA-----ACTGAC 131
 146 GAGATTAAGTGGAGCTGTGCAAGCGCTGCGCTTGGCGCGGAGAGCGCGCTGACCA 225
 132 CAGATTAAGTGGAGCTGTGCAAGCGCTGCGCTTGGCGCGGAGAGCGCTGACCA 191
 206 ACCGATTAAGTGGAGCTGTGCAAGCGCTGCGCTTGGCGCGGAGAGCGCTGACCA 265
 192 GCTTGTCTTAAGTGGAGCTGTGCAAGCGCTTGGCGCGGAGAGCGCTGACCA 251
 256 CTGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 325
 250 CTGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
 326 TTTGCTTAAGTGGAGCTGTGCAAGCGCTGCGCGGAGAGCGCGCTGACCA 385
 312 TGTGCTTAAGTGGAGCTGTGCAAGCGCTGCGCGGAGAGCGCGCTGACCA 371
 386 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 445
 372 CCGATTAAGTGGAGCTGTGCAAGCGCTGCGCGGAGAGCGCGCTGACCA 431
 446 GTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 505
 432 GTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 491
 506 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565
 492 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 551
 566 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 625
 552 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611
 626 CCAAGTGGGCTTGAAGTGTGCAAGAGCGGCAAGCGGCAAGTGTGCAAGAGT 685
 612 CCAAGTGGGCTTGAAGTGTGCAAGAGCGGCAAGCGGCAAGTGTGCAAGAGT 671
 686 GCAAGTGGGCTTGAAGTGTGCAAGAGCGGCAAGCGGCAAGTGTGCAAGAGT 745
 672 GCAAGTGGGCTTGAAGTGTGCAAGAGCGGCAAGCGGCAAGTGTGCAAGAGT 731
 746 CCGCTGCTTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803
 732 CCGCTGCTTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 769

RESULT 4
 AAV15072
 ID AAV15072 standard; DNA; 922 BP.

XX AAV15072:
 AC 17-OCT-2003 (revised)
 XX 19-MAY-1998 (first entry)
 DT Hybrid DNA comprising a family 45 cellulase core region.
 XX Cellulase core region; isolation; microorganism; identification;
 XX hybrid DNA; ss.
 OS Humicola nigrescens.
 OS Humicola insolens.
 OS Chimeric.
 XX Key Location/Qualifiers
 FT 1..921
 FT CDS /*tag= a
 FT /product= "hybrid gene construct"
 XX MO9743409-A2.
 XX 20-NOV-1997.
 XX 12-MAY-1997; 97MO-DK00216.
 XX 10-MAY-1996; 96DK-00000562.
 XX (NOVO) NOVO-NORDISK AS.
 XX Dalboge H, Diderichsen B, Sandal T, Kauppinen S;
 XX MPI, 1998-008878/01.
 XX P-PsDB: AAM44259.
 PT Isolating novel DNA sequences from microorganisms - without the need for
 PT culturing the microorganism.
 XX Example 3; Page 51-52; 72pp; English.
 CC The present sequence represents a novel hybrid gene construct from an
 CC example of the present invention. The present invention describes a novel
 CC method for providing a novel DNA sequence encoding a polypeptide from a
 CC microorganism with an activity of interest. The method comprises: (i) PCR
 CC amplification of the DNA with PCR primers with homology to (a) known
 CC gene(s) encoding a polypeptide with an activity of interest; (ii) linking
 CC the obtained PCR product of a 5' structural gene sequence and a 3'
 CC structural gene sequence; (iii) expressing the resulting hybrid DNA
 CC sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide
 CC with the activity of interest or a related activity; and (v) isolating
 CC the hybrid DNA sequence identified in step (iv). This method provides for
 CC identification and isolation of sequences from microorganisms without
 CC having to cultivate and isolate the microorganism. (Updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 XX Sequence 922 BP; 151 A; 333 C; 262 G; 176 T; 0 U; 0 Other;

Query Match 39.5%; Score 451.6; DB 2; Length 922;

Best Local Similarity 71.6%; Pred. No. 8.5e-73;

Matches 630; Conservative 0; Mismatches 229; Indels 21; Gaps 2;

102 TTCTCTTGGTGGCTCTGCGGCGGCAAGTGGCAAGTGTGCAAGATTAAGTGGAGTGG 161
 43 CTTCAAGTGGCGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 102
 162 TCGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 221
 103 TGTAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 162
 222 GATGCAAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 281
 163 AAGCAAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 222

QY	282	GCCAACTCCCTGCGCGACACAGACTCCCTGCGCGGTGAAGACAATCTCGCTTAACGGCTTC	341
Db	223	GCTCTCTCTGCGCGCGACACAGACCCCGTGGCGGTGAAGACAATCTCGCTTAACGGCTTC	282
QY	342	GCGCGACACAGAGATCGCCGCGCGGTTCGGAATCTCTGCTGGTGGTGGCTGAACCGCTC	401
Db	283	GCGGCTACCGGCGCTCGCGCGCGCAGTCCGAGCTTCTGGTGGTGGTGGCTGAACCGACTC	342
QY	402	ACCTTCACCTTCGCGTCCCGTCCCGCGCGACAGACAATGGTGGTGGCTGAACAGACCTGGC	461
Db	343	ACCTTCACCTTCGCGTCCCGTCCCGCGCGACAGACAATGGTGGTGGCTGAACAGACCTGGC	402
QY	462	GACGACCTCGGAAAGTACACAGTTCGATATCGCATGCGCGCGCGCGGTGGGCAATCTTC	521
Db	403	GATGACCTCGGAAAGTACACAGTTCGATATCGCATGCGCGCGCGGTGGGCAATCTTC	462
QY	522	AAACGCTGCAAGTCTGACGATTGGCGCGCTCCCGCGCGCTCAATACCGCGCGCAATTCGTG	581
Db	463	GACCGCTGCTCCCTCGAGTGGCGGGTCTCCCGCGCGCAACGCGCTATGGCGCGCGCTGCTC	522
QY	582	CGCGACCAAGTGGCAATCTCTCCCGCGCGCGCGTCAACCGCGCGCGGTGGGCAATCTTC	641
Db	523	CGCGACCAAGTGGCAATCTCTCTCCCGCGCGCGTCAACCGCGCGCGGTGGGCAATCTTC	582
QY	642	TGGCTTCAGAAACGCGCGACAACCGCGAGTTCACGTTCCAGACAGTGGAGTGGCGCGCGAG	701
Db	583	TGGCTTCAGAAACGCGCGACAATCCGAGCTTCAGTTCGATCGAGTTCAGTGGCGCGCGAG	642
QY	702	ATCGTTGCGCGCTCGCTCGCGCTCGAAAGGCGACAGACATCTCGAGCTTCCCGGTCTTACACCCC	761
Db	643	CTGCTGCTCGCGCGCGAGTCCGCGCGCGACAACAGCGCGCAATCTCCCTGCGCGCGAGTC	702
QY	762	CCAAAGCGATGGCAACCG-----TGGCACCGGGACGCGCCAGTGGACT	803
Db	703	CCCTTCAGAGACACACACTCTCGGTGACACAGCTACAGACACAGACACAGTTCACCC	762
QY	804	GCGCGTGGGTGGGTCAGAC--GTCTCCCGCGCGCGCGCGAGTGGCTGACAGTTCAGAG	860
Db	763	TCCACCACTCGAGCGCGCGCGAGTTCAGACCTTACAGACTCCACGCGGCTGCACTGCTGAGAG	822
QY	861	TGGGCTCAGTGGGTGGTGCATCGGCTTCAGCGGAGTGAACACACCTGCTGCTGGACACAC	920
Db	823	TGGGCTCAGTGGGTGGGGAATGGCTGAGCGCGGAGACACACCTGGTGGTGGGACACACT	862
QY	921	TGCCAGAGTTGAACGACACTACTCTGCGAGTGCCTCTAA	960
Db	883	TGCACGAGATTATTAATGACTGTACATCAGTGGCTGTAA	922
RESULT 5			
AAZ60178 standard; cDNA; 1060 BP.			
AC	AAZ60178;		
DT	25-APR-2000	(first entry)	
XX	H. insoliens	endoglucanase enzyme coding sequence.	
XX	Endoglucanase; cellulase enzyme; detergent composition; anion surfactant;		
XX	degradation protection; liquid detergent; long term stability; ss.		
OS	Hemicola insolens.		
XX	EP6333-i-1.		
PN	11-JAN-1995.		
PD	28-JUN-1993;	93EP-00870122.	
PF	28-JUN-1993;	93EP-00870122.	
PR	(PROC) PROCTER & GAMBLE CO.		

XX		Hierboots IMJ, Jansen MP;
XI		MFI; 1995-038508/06.
XX		P-PSDB; AAY67388.
DR		
XX		Liq. detergent compens. - consg. hydrophobic amine(s) for cellulase stabilisation.
PT		
XX		Disclosure; Page 16-17; 28pp; English.
CC		This sequence represents the Humicola insolens endoglucanase enzyme coding sequence. The enzyme is a cellulase used in the liquid detergent composition of the invention. The detergent comprises an anion surfactant, a cellulase enzyme and a cellulase stabilising amount of an amine of the formula R ₁ R ₂ NH ₂ , where R ₁ and R ₂ are independently H or a C2-C9 alkyl chain, and N is a C2-C9 alkyl chain or cyclohexyl or cyclopentyl or cycloheptyl. The amine prevents the cellulase enzyme from becoming degraded. The composition is used as a liquid detergent, and has long term stability; it is applicable to the protection of an cellulase and can be used in the presence of proteases
CC		
XX		
SQ	Sequence 1060 BP; 190 A; 378 G; 287 G; 205 T; 0 U; 0 Other;	
	Query Match	37.4%; Score 439; DB 2; Length 1060;
	Best Local Similarity	70.5%; Pred. No. 1.7e-70;
	Matches 656; Conservative	0; Mismatches 245; Indels 30; Gaps 4;
OY	55 CGAGATGCGCTACTACCTCCGGTCTTGTCGCACAACCCTGCGGTCACTTTCTGTGGT	114
Dn	5 CCAGAATGGGTCTCCTCCCCCTCCCGGTGCGCGGTGAGGCCGCTGCGGTGTGG 64	
OY	115 CCTCCCGCGCACGATGACGATGACGACGACGAATACTGGAACTTGTCGCAAGCGCTGT	174
Dn	65 CCTTCGCGCATGGACAG-----GTCCACCOCGTACTGGAACTGTGCAAGCTTTGT	118
OY	175 GCCTTTGACCGGAAAAGCCGCTCAAGCCAACCGATTCAAAGCTGCAATGCCATTGCC	234
Dn	119 GGCGCTGGCCCAAAAGAGCTCCCCTGGAACAGCTGTCTTTTCTGCAACGCCAATTCC	178
OY	235 AGGCGCTGTGCACTTCATATGTCACAGTCCGAGCGCA---CGCGGTCCGCTCACTCC	291
Dn	179 AGGGTAACAGAACTTAGACGCAAGTCCGAGTCGACACCGCGGAGTGTCCGTACTGT	238
OY	292 GCGCGACACGACTCCCTGGCGGTGACAGAACAACTATCGCTCAAGCTTGCDCGACGA	351
Dn	239 GCCCGACCAACCCCACTGGGTGGAAGCAAGACTTGCGGTGGTTTTCTGTCACT	298
OY	352 GCATGCTGCGGAGGTCCGAATCTTCGTGAGTGGCTGTGTAAGGCTTCACTTCACTT	411
Dn	299 CTATTCCCGGAGAAATGAGAGCGAGCTGTGCTGTGCTGTCAAGCTCACTTCACT	358
OY	412 CCGGTCCGCTGCGCGCAGACCAATGATGATGTCATGCAAGAACATGGCGACGACTCG	471
Dn	359 CCGGTCTGTTGTGCGGAAGAAATGATGTCTGTCTGATCAACAATCGGCGATGATCTTG	418
OY	472 GAATGACCAAGTTGATATGTCATGTCGCGCGGCGGCGGTGGGCACTTTCAAGCGCTGA	531
Dn	419 GCGAACAACATTTGATCTCAATATCCCGCGAGCGGTGTGGCAATCTTCAAGCAATGA	478
OY	532 GCTGCAAGTTGAGGAGCTCCCGCGGCGCTCAATTAAGAGGCAATTTCGTGCGGCAACAT	591
Dn	479 CTCCTCAAGTTGAGGAGCTCTGCGCGGCGAGCGCTCAAGCGGCAATCTGTCCGCAAGAT	538
OY	592 GCGATTTCTTCCCGCGCGCTCAAGCGCGGCTGCGCAAGTGGAGTGTGATCTGTGTCAGA	651
Dn	539 GCGATGGTTTCCCGCGAGCCCTCAAGCGCGGCTGTCACTGTGCTTCACTGTGTTCAAA	598
OY	652 AACCGCAACAACCGAAGTTCAAGTTTCAAGAGTGGCAAGTCCCGCGCAAGTGTGTGCC	711
Dn	599 AACCGCAACAATTCGAAGTTCAAGTTTCCGTGATGCAATGCGCAAGCGTGTGTGCTGTC	658
OY	712 GCTCGCGCTGCAAGCGCAAGCAAGTCAAGTCTCCCGCTTTCACACCCCCCAAGCGGTG	771


```

XX AC AAQ14856;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 18-FEB-1993 (first entry)
XX XX
XX DE Humicola insolens DSM 1800 endoglucanase.
XX XX
XX KW cellulase; ss.
XX XX
XX OS Humicola insolens.
XX XX
XX FH Key Location/Qualifiers
XX FT sig_peptide 10..72
XX FT /*tag= a
XX FT mat_peptide 73..924
XX FT /*tag= b
XX XX
XX EN W09117243-A.
XX XX
XX PD 14-NOV-1991.
XX XX
XX PF 09-MAY-1990; 90DK-00001159.
XX XX
XX PR 09-MAY-1990; 90DK-00001159.
XX PR 22-APR-1991; 91DK-00000736.
XX XX
XX PA (NOVO ) NOVO-NORDISK AS.
XX XX
XX PI Rasmussen G., Mikkelsen JW, Schulein W, Packar SA, Hagen F;
XX PI Hjort C, Hestrup S;
XX DR WP1; 1991-353/65/48.
XX DR P-PSDB; AAR15271.
XX XX
XX PT Cellulase prep. comprising endoglucanase enzyme - used in detergents for
XX PT cellulose-contg. fabrics or to improve drainage of paper pulp.
XX PS Claim 11; Page 48; 67pp; English.
XX XX
XX CC The cellulase (i.e. endoglucanase) gene was isolated from a H. insolens
XX CC cDNA library by screening with a probe based on the amino acid sequence
XX CC of the purified enzyme. (updated on 25-MAR-2003 to correct PA field.)
XX CC
XX SQ Sequence 1060 BP; 190 A; 377 C; 288 G; 205 T; 0 U; 0 Other;
XX XX
XX Query Match 37.3%; Score 437.4; DB 2; Length 1060;
XX Best Local Similarity 70.4%; Pred. No. 3.3e-70;
XX Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;
XX XX
XX QY 55 CCAAGATGCGCTTACTCCGTTCTTCGCAACCTGCGCGCTGCACTTCTCTGTGTG 114
XX DB 5 CCAAGATGCGTTCTCCCTCCCTCTCCGTCGCGCTGTGTGCGCGCTGCGTGTGTG 64
XX QY 115 CTTCCCGGCGCATGTGCGCATTCACAGATCTGCGCATCTGTGCAAGCGCTGCT 174
XX DB 65 CCTTCGCGCATGTGCGAG-----GTTCACCCCTACTGCGCATCTGTGCAAGCTTGTG 118
XX QY 175 GCGCTTGCGCGGGAAGGCGCGCTGCGCAACCGCTGCTGCGCATCTGCTTCTCC 234
XX DB 119 GCGCTTGCGCGCAAGAGCTCCCTGCAACCGCTGCTTCTGCAAGCGCAACTTCC 178
XX QY 235 AGCGCTGTGCGCATCTGCAATGTCAGTGGGTGCA--CGCGCGCTGCGCTTCTCT 291
XX DB 179 AGCGTATGACGAGCTTGCAGCGCAAGTCCGCTCGAGCGCGCGCTGTGCTTACTCGT 238
XX QY 292 GCGCGCAAGCATCTCCCTGCGCGGTGAGACAGCATCTGCTAGCGCTTCCGCGCA 351
XX DB 239 GCGCGCAAGCATCTCCCTGCGCGGTGAGACAGCATCTGCTAGCGCTTCCGCGCA 298
XX QY 352 GCATTCGCGCGCGGTGCGCAATCTCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCT 411
XX DB 299 CTATTCGCGCGCAAGCATCTGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

```

```

QY 412 CCGGTCCCTGCGCGGCAAGCAATGTGTGTGCTGCTGCAAGCACTGCGCGCACTGCG 471
DB 359 CCGGTCTCTGTGTGCTGCAAGCAAGTGTGTGTGCTGCTGCAAGCACTGCGCGCACTGCG 418
QY 472 GAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
DB 419 GCAAGCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478
QY 532 GCTGCAAGTGTGCGCGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
DB 479 CTCCCAAGTGTGCGCGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
QY 592 GCGATTCCTTCCCGCGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
DB 533 GCGATTCCTTCCCGCGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
QY 652 AGCGCGCAAGCAAGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711
DB 599 AGCGCGCAAGCAAGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
QY 712 GCTGCGCTGCAAGCGCAAGCAAGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
DB 659 GCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
QY 772 GCAAGCG-----TGGACCGCGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
DB 719 GCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
QY 814 CCGCGCAAGC---GTCTCCCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
DB 779 CGAGCGCGCGCAAGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 838
QY 871 GCGGTGCGCATGCGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
DB 839 GCGGTGCGCATGCGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898
QY 931 TGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
DB 899 TTAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 929

RESULT 8
AAQ25932
ID AAQ25932 standard; DNA; 1060 BP.
XX AC AAQ25932;
XX XX
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 15-JUN-1993 (first entry)
XX XX
XX DE Cellulase contained in a detergent compsn.
XX XX
XX KW Endoglucanase; immunoreactive; Humicola insolens; cleaning; softening; ss.
XX XX
XX OS Humicola insolens.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 9..927
XX FT /*tag= a
XX XX
XX EN E495554-A1.
XX XX
XX PD 22-JUL-1992.
XX XX
XX PF 15-JAN-1992; 92EP-00200101.
XX XX
XX PR 16-JAN-1991; 91EP-00870006.
XX PR 06-NOV-1991; 91EP-00202881.
XX XX
XX PA (PROC ) PROCTER & GAMBLE CO.
XX XX

```


Db 659 GCACCGATGCGCGCCGACAGACGAGCACTTCCTCCGTCCAGATCCCTCCAGCA 718
 QY 772 GCACCG-----TGGCACCGGAGCGCCACCTGTGAGCTGCGCTGGGT 813
 Db 719 GCACCGATGCTCCGCTCAACGACGCTTACGACGACGACGACGCTTCCAGCTCCAGCACT 778
 QY 814 CGGCGCAGAC---GTCTCCGCGCGCGGAGTGGCTGACAGCTTCAAGATGGCTTCAGT 870
 Db 779 CGAGCCGCGCAGTCCAGCTTACGACTCCGCGGCTGACCTGCTGAGAGGTGGCTCAGT 838
 QY 871 GCGGTGATGGCTTACGCGGATGACGACCACTGTGTCTTGGACACCACTGCGCAGAGT 930
 Db 839 GCGGCGGATGCTGGAGCTGCTGACGACGACCTGCGTGGCTGGACGACTTGCAGAGAA 898
 QY 931 TGAACGACTACTACTGCGAGTGCCTTAAAC 961
 Db 899 TTAATGACTGTACCAATGAGTGCCTGTAAG 929

RESULT 12

AAQ41732
 ID AAQ41732 standard; DNA; 1060 BP.

AC AAQ41732;

DT 25-MAR-2003 (revised)
 DT 25-AUG-1993 (first entry)

DE Dye transfer inhibiting compsn. cellulase gene.

XX Detergent; homogeneous endoglucanase component; 43kd cellulase; ss.

CS Humicola insolens.

Key Location/Qualifiers

FT CDS 10..927

FT sig_peptide /*tag= a

FT mat_peptide /*tag= b

FT /*tag= c

PN EP540784-AL.

PD 12-MAY-1993.

PF 06-NOV-1991; 91EP-00202882.

PR 06-NOV-1991; 91EP-00202882.

PA (PROC) PROCTER & GAMBLE CO.

PI Mccorquodale F, Busch A;

DR WPI; 1993-153868/19.

DR P-PSDB; AAR37150.

PT Compn. for inhibiting dye transfer during fabric washing - contains peroxidase, hydrogen peroxide, substrate and cellulase, esp. endo-glucanase from Humicola insolens.

PS Disclosure; Page 18-19; 28pp; English.

CC The sequence is that encoding a cellulase which is characterised in that the cellulase provides at least 10% removal of immobilised radioactive CC labelled carboxymethyl cellulose according to the CIAWC method at 25 x CC 10(-6) by weight of the cellulase protein in the test solution. It can be used as part of a compsn. for inhibiting dye transfer. [Updated on 25-MAR-2003 to correct PN field.]

XX Sequence 1060 BP; 190 A; 377 C; 288 G; 205 T; 0 U; 0 Other;

Query Match 37.3%; Score 437.4; DB 2; Length 1060;

Best Local Similarity 70.4%; Pred. No. 3.3e-70;
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CCAGATGCGCTTACTCTCCGTTCTTTCGACACCTTGGCGCTGACCTTCTCTGTGG 114
 Db 5 CCAAGATGGTTCCTCCCTCCCTCTCCGTCGCGCTGTGTGGCGCGCTCGCGGTGG 64
 QY 115 CTTCCGCGCAGATGGCAATGGGCAATGCTCAAGATATCTGGAGCTCTGCAACCGCTGT 174
 Db 65 CCTTCGCGCTGATGGCAG-----GTCCACCGCTACTGGAGCTCTGCAAGCGCTTGT 118
 QY 175 GCGCTTGGCGCGGAGAGCGCGCTGACGACCGGCTTCAAGGCTGAGTGGCAATCTCC 234
 Db 119 GCGCTGGCGCAAGAGCTCTCCGTAACACCGCTGTCTTCTTGGCAAGCGCACTTCC 178
 QY 235 AGCGCTGTCCGACTTCAATGTCCAGTCCGGCTGCA---CGCGGCTGTGCTACTTCT 291
 Db 179 AGCGTATCAAGAGCTTCCAGCGCAATCGGCTGTGGAGCGCGGCTGTGCTACTTCT 238
 QY 292 GCGCGCAGCACTCTCTGGCGGTGAACGACATCTGCTTACGGCTTGGCGCGACGA 351
 Db 239 GCGCGCAGCAACCCCATGGGCTGTGACGACGACTTGGGCTGTGCTTGTGCTGCACT 298
 QY 352 GCATGCGCGGCGGAGTCCGAACTCTGAGTGGTGGCTGCTGCTACAGCGCTCACTTCACT 411
 Db 299 CTATGCGCGGAGCAATGAGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
 QY 412 CCGGCTCCGCTGCGCGGCAAGCAATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 471
 Db 359 CCGGCTCTGTGCTGCGCAAGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418
 QY 472 GAGGTACCAAGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
 Db 419 GCGACCAACCACTTGTGATCTCAATCTCCCGGCGGCGGCTGCTGCTGCTGCTGCTG 478
 QY 532 GCTGCGAGTTGGGCGGCTCCCGCGGCTCAATGAGGCGGCTTGTGTGCGCGACAGT 591
 Db 479 CTCCCGAGTTGGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 538
 QY 592 GCGATTCCTTCCCGCGCGCTGCAAGCGGCTGCGGCTGCGGCTGCTGCTGCTGCTG 651
 Db 539 GCGATCGGTTCCCGAGCGCTTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
 QY 652 AGCGCGACCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
 Db 599 AGCGCGACCAATCCGAGCTTGAAGCTTCCGTCAGTCCGCTGCTGCTGCTGCTGCT 658
 QY 712 GCTCCGCGCTGCAAGGCGCAAGAGCTTCCGCTTCCGCTTCAACCCCGCAAGCGGTG 771
 Db 659 GCACCGGATGCGCGCGGCAAGAGGCGCACTTCCGCTGCTGCTGCTGCTGCTGCTG 718
 QY 772 GCACCG-----TGGCACCGGAGCGCCGACGCTGCACTGCGCTTGGGT 813
 Db 719 GCACCGATGCTCCGCTCAACGACCTTACGACGACGACGACGACGCTTCCAGCACT 778
 QY 814 CGGCGCAGAC---GTCTCCGCGCGCGGAGTGGCTGCACTTCCAGAGTGGCTCACT 870
 Db 779 CGAGCCGCGCAGTCCAGCTTACGACTCCAGCGGCTGCACTTGGAGGTTGGCTCACT 838
 QY 871 GCGGTGATGCGGCTTACGCGGATGACCACTGTGTCTTGGACACCACTGCGCAGAGT 930
 Db 839 GCGGCGGCAATGCTGAGAGGCGCTGCAACCACTGTGTGCTGCGACACTTGCAGAGAA 898
 QY 931 TGAACGACTACTACTGCGAGTGCCTTAAAC 961
 Db 899 TTAATGACTGTACCAATGAGTGCCTGTAAG 929

RESULT 13

AA110182
 ID AA110182 standard; DNA; 1060 BP.

AC AA110182;

XX MO9743409-A2.
 EN 20-NOV-1997.
 PD 12-MAY-1997; 97WO-DK000216.
 EF 10-MAY-1996; 96DK-00000562.
 PR (NOVO) NOVO-NORDISK AS.
 XX Dalboge H, Diderichsen B, Sandal T, Kauppinen S;
 PI WPI; 1998-008878/01.
 DR P-PDB; AAM44266.
 XX Isolating novel DNA sequences from microorganisms - without the need for
 PT culturing the microorganism.
 XX Example 3; Page 47-48; 72pp; English.
 PS The present sequence encodes a protein with endoglucanase activity used
 CC in an example of the present invention. The present invention describes a
 CC novel method for providing a novel DNA sequence encoding a polypeptide
 CC from a microorganism with an activity of interest. The method comprises:
 CC (i) PCR amplification of the DNA with PCR primers with homology to (a)
 CC known gene(s) encoding a polypeptide with an activity of interest; (ii)
 CC linking the obtained PCR product of a 5' structural gene sequence and a
 CC 3' structural gene sequence; (iii) expressing the resulting hybrid DNA
 CC sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide
 CC with the activity of interest or a related activity; and (v) isolating
 CC the hybrid DNA sequence identified in step (iv). This method provides for
 CC identification and isolation of sequences from microorganisms without
 CC having to cultivate and isolate the microorganism
 XX
 SQ Sequence 1060 BP; 190 A; 377 C; 288 G; 205 T; 0 U; 0 Other;
 Query Match 37.3%; Score 437.4; DB 2; Length 1060;
 Best Local Similarity 70.4%; Pred. No. 3.3e-70;
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 532 GCTGCGAGTTGCGCGCGCTCCCGCGCTCAATACGCGCATTTGTCGCGACAGT 591
 DB 479 CTCGCCAGTTGCGCGCTGTCGCGCGACGCGTACGCGCGCATCTGTCGCGCAAGAGT 538
 QY 592 GCGATTCTTCCCGCGCGCTGTAAGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 651
 DB 539 GCGATCGGTTCCCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 598
 QY 652 ACGCGCAACCCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 711
 DB 599 ACGCGCAACCCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 658
 QY 712 GCTCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 771
 DB 659 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 718
 QY 772 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 813
 DB 719 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 778
 QY 814 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 870
 DB 779 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 838
 QY 871 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 930
 DB 839 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 898
 QY 931 GCGCGCGTTCGCGCGCGCTGTAAGCGCGCTGCGCGCTGCGCGCTTTGATCTGTTCCAGA 961
 DB 899 TTAATGACTGTAACATGATGCTGTAAGC 929

RESULT 15
 ID AAV16102 standard; cDNA; 1060 BP.
 XX AAV16102;
 AC 21-JUL-1998 (first entry)
 XX
 DT Humicola insolens surface-active endoglucanase gene.
 DE
 XX
 KW endoglucanase; surface-active; cellulase; detergent; cleaning;
 XX performance; stain removal; soften; feel; colour; ss.
 XX
 OS Humicola insolens.
 XX
 FH Key Location/Qualifiers
 FH CDS 10..927
 FT /*tag= a
 FT /product= "endoglucanase"
 FT sig_peptide 10..72
 FT /*tag= c
 FT mat_peptide 73..924
 FT /*tag= b
 XX
 PK WO9804663-A1.
 XX
 PD 05-FEB-1998.
 XX
 PF 25-JUL-1997; 97WO-US013194.
 XX
 PR 30-JUL-1996; 96US-0023125P.
 XX
 PA (PROC) PROCTER & GAMBLE CO.
 XX
 PI Boyer SL;
 XX
 DR MPI; 1998-130664/12.
 DR P-PDB; AAM46616.
 XX
 PT Detergent composition containing both surface-active and non-surface

PT active cellulase - softens and improves feel of cotton fabrics without
PT causing loss in weight or tensile strength.

XX Disclosure; Page 41-43; 68pp; English.

XX The sequence is that encoding the endoglucanase enzyme of a surface-
CC active cellulase produced by Humicola. The enzyme can be used in a
CC detergent composition with a non-surface-active cellulase. This
CC combination improves cleaning performance (maintains colours and removes
CC stains), and softens and improves the feel of cotton fabrics without
CC causing losses in weight or tensile strength

XX Sequence 1060 BP; 190 A; 377 C; 288 G; 205 T; 0 U; 0 Other;

Query Match 37.3%; Score 437.4; DB 2; Length 1060;

Best Local Similarity 70.4%; Pred. No. 3.3e-70;

Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

```
QY 55 CCAAGATGCGCTCTACTCCGTTTGTGCAACCCCTGCGCTGCACTTCTCTGCTG 114
DB 5 CCAAGATGCGCTCTACTCCGTTTGTGCAACCCCTGCGCTGCACTTCTCTGCTG 64
QY 115 CCGTCCGCGCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 174
DB 65 CCGTCCGCGCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 118
QY 175 GCGCTTGGCCCGGGAAGCGCGCGCTGAGCGCAACCGGCTTACGCGTGAAGCACTTCC 234
DB 119 GCGCTTGGCCCGGGAAGCGCGCGCTGAGCGCAACCGGCTTCTTCTTCTGCAAGCGCACTTCC 178
QY 235 AGCGCTGTGCGCACTTCAATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 291
DB 179 AGCGCTGTGCGCACTTCAATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 238
QY 292 GCGCGCAACGAGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 351
DB 239 GCGCGCAACGAGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 298
QY 352 GCGCGCAACGAGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 411
DB 299 GCGCGCAACGAGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 358
QY 412 CCGGTCCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 471
DB 359 CCGGTCCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 418
QY 472 GAGTAAACGAGTTCGATTCGCAATGCGCGGCGGCGGCGGCGGCGGCGGCGG 531
DB 419 GAGTAAACGAGTTCGATTCGCAATGCGCGGCGGCGGCGGCGGCGGCGGCGG 478
QY 532 GCTCGAGTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 591
DB 479 CTCCCGAGTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 538
QY 592 GCGATTCCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 651
DB 539 GCGATTCCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 598
QY 652 AGCGCGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 711
DB 599 AGCGCGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 658
QY 712 GCTCCGCGTTCGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 771
DB 659 GCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 718
QY 772 GCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 813
DB 719 GCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 778
QY 814 GCGCGCAACGAGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 870
DB 779 CGAGCGCGCGAGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 838
```

```
QY 871 GCGGCGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 930
DB 839 GCGGCGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 898
QY 931 TGAACGAGTTCGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 961
DB 899 TGAACGAGTTCGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 929
```

Search completed: July 7, 2004, 08:15:27
Job time : 723 secs


```

Qy 541 TCGGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600
Db 541 TCGGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600
Qy 601 TCGGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600
Db 601 TCGGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600
Qy 661 ACCGAGCTTCAAGTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 720
Db 661 ACCGAGCTTCAAGTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 720
Qy 721 GCAAGCGCAAGCAAGCTCAAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 780
Db 721 GCAAGCGCAAGCAAGCTCAAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 780
Qy 781 GCACCGGAGCGGCGGCTCAAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 840
Db 781 GCACCGGAGCGGCGGCTCAAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 840
Qy 841 GTGCTGCTCAAGTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 900
Db 841 GTGCTGCTCAAGTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 900
Qy 901 CCGTGTCTCTGCAAGCTCAAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 960
Db 901 CCGTGTCTCTGCAAGCTCAAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 960
Qy 961 CAGCTTTTCCGCAAGTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 1020
Db 961 CAGCTTTTCCGCAAGTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 1020
Qy 1021 TTTTGAAGGCTCAATAGTTCAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 1080
Db 1021 TTTTGAAGGCTCAATAGTTCAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 1080
Qy 1081 TCTCAAGCGCAAGCTTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 1140
Db 1081 TCTCAAGCGCAAGCTTTCAGAGGTCAGTGCCTCCCGGAGATGTTGCGCGCT 1140
Qy 1141 TTTTGAAGGCTCAATAGTTCAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 1174
Db 1141 TTTTGAAGGCTCAATAGTTCAGTTCCTCCCGCTTTCACCCCGGAGCGTGC 1174

```

RESULT 3

US-09-229-911A-11

Sequence 11, Application US/09229911A

Patent No. 6387690

GENERAL INFORMATION:

APPLICANT: Schuelein, Martin

Andersen, Lene N.

Lassen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Rudy I.

Ihara, Michiko

Takagi, Shinobu

TITLE OF INVENTION: No. 6387690e1 Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESS: No. 6387690 No. 6387690disk of No. 6387690th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/229, 911A

FILING DATE: 13-Jan-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366-200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 1174 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 60..956

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-229-911A-11

Query Match 100.0%; Score 1174; DB 4; Length 1174;

Best Local Similarity 100.0%; Pred. No. 1.9e-225;

Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GAGGAGACCCCTTCAAGCTGTAAGTTTCACCCCGCTCTTTTCTTGGCCCCGAGA 60
Db 1 GAGGAGACCCCTTCAAGCTGTAAGTTTCACCCCGCTCTTTTCTTGGCCCCGAGA 60
Qy 61 TGGGCTCTACCTCGGTTCTTTCGCAACAACCTGCGCTGCACTTCTCTGTCGCTCCG 120
Db 61 TGGGCTCTACCTCGGTTCTTTCGCAACAACCTGCGCTGCACTTCTCTGTCGCTCCG 120
Qy 121 CGGCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 CGGCAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Qy 181 GGGCGGGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 GGGCGGGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Qy 241 TGTCCGACTTCAATGTCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAGGCT 300
Db 241 TGTCCGACTTCAATGTCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAGGCT 300
Qy 301 AGACTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 301 AGACTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Qy 361 GCGGCTCCGATCTCTGAGTGCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAG 420
Db 361 GCGGCTCCGATCTCTGAGTGCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAG 420
Qy 421 TCGCGGCAAGCAAGTTCGAGTGCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAG 480
Db 421 TCGCGGCAAGCAAGTTCGAGTGCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAG 480
Qy 481 AGTTCGATTCGCAAGTTCGAGTGCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAG 540
Db 481 AGTTCGATTCGCAAGTTCGAGTGCAGTTCGAGCTGCAAGCGGCTGAGGCTGAGGCTGAG 540
Qy 541 TCGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600
Db 541 TCGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600
Qy 601 TCGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600
Db 601 TCGCGGCTCCCGGCTCAATAGGCGGATTTGCGCGCAAGTCGATTCCT 600

```


COUNTRY: USA
ZIP: 45217
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/090,013
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 9120280.0
FILING DATE: 06-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: PATEL, KEN K.
REGISTRATION NUMBER: 33,988
REFERENCE/DOCKET NUMBER: CM393
TELECOMMUNICATION INFORMATION:
TELEPHONE: 513-627-6437
TELEFAX: 513-627-4854
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1060 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 10..924
US-08-090-013-1

Query Match 37.3%; Score 437.4; DB 1; Length 1060;

Best Local Similarity 79.4%; Pred. No. 1.3e-78;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

55 CCAAGATGCGCTCTTACCTCCGTTCTTGGCAACCCCTGCGCTGCACTTCTCTGCTG 114
5 CCAAGATGCGCTCTTACCTCCGTTCTTGGCAACCCCTGCGCTGCACTTCTCTGCTG 64
115 CCTCGCGCGCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 174
65 CCTCGCGCGCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 118
175 GCGCTGCG 234
119 GCGCTGCG 178
235 AGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCT 291
179 AGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCT 238
292 GCGCGCGCGCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 351
239 GCGCGCGCGCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 298
352 GCG 411
299 CATTGCG 358
412 CCGCTGCG 471
359 CCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCT 418
472 GAAGTACCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCT 531
419 GCGCGCGCGCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 478
532 GCG 591
479 CTTCCGAGTTTGGGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 538
592 GCGATTCTTCCCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCA 651

RESULT 6
US-08-081-328-1
Sequence 1, Application US/08081328

Patent No. 5520838
GENERAL INFORMATION:

APPLICANT: BARCK, ANDRE C.
APPLICANT: BUSCH, ALFRED (NMN)

TITLE OF INVENTION: COMPACT DETERGENT COMPOSITIONS WITH
TITLE OF INVENTION: HIGH ACTIVITY CELLULOSE

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: THE PROCTER & GAMBLE COMPANY
STREET: 11810 East Miami River Road

CITY: CINCINNATI
STATE: OHIO

COUNTRY: USA
ZIP: 45253-8707

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/081,328
FILING DATE: 11/19/93

CLASSIFICATION: 252
ATTORNEY/AGENT INFORMATION:

NAME: ZERBY, KIM WILLIAM
REGISTRATION NUMBER: 32,323

REFERENCE/DOCKET NUMBER: CM356M
TELECOMMUNICATION INFORMATION:

TELEPHONE: 513-627-2885
TELEFAX: 513-627-0318

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1060 base pairs
TYPE: nucleic acid

STRANDEDNESS: unknown
TOPOLOGY: unknown

FEATURE:
NAME/KEY: CDS

LOCATION: 10..924
US-08-081-328-1

DB 539 GCGATCGCTTCCCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 598
QY 652 ACCCGCGCAACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 711
DB 599 ACCCGCGCAACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 658
QY 712 GCG 771
DB 659 GCG 718
QY 772 GCG 813
DB 719 GCG 778
QY 814 GCG 870
DB 779 GCG 838
QY 871 GCG 930
DB 839 GCG 898
QY 931 TGAACGCTACTGCG 961
DB 899 TTAATGACTGTACCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 929

Query Match 37.3%; Score 437.4; DB 1; Length 1060;
Best Local Similarity 70.4%; Pred. No. 1.3e-78;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 33 CAGAGATGCGCTCTATCTCCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGAG 114
DB 5 CCAAGATGCGCTCTATCTCCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGAG 64
QY 115 CTTCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 174
DB 65 CTTCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 118
QY 175 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 234
DB 119 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 178
QY 235 AGCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 291
DB 179 AGCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 238
QY 292 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 351
DB 239 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 298
QY 352 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 411
DB 299 CTTCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 358
QY 412 CCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 471
DB 359 CCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 418
QY 472 GAGTAACTGCGCTTCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 531
DB 419 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 478
QY 532 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 591
DB 479 CTTCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 538
QY 592 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 651
DB 539 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 598
QY 652 AGCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 711
DB 599 AGCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 658
QY 712 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 771
DB 659 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 718
QY 772 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 833
DB 719 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 778
QY 834 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 870
DB 779 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 838
QY 872 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 930
DB 839 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 896
QY 931 TGAAGCTTCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 961
DB 899 TGAAGCTTCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 929

RESULT 7

US-08-232-249-1

; Sequence 1, Application US/08232249

; Patent No. 5610129

GENERAL INFORMATION:
APPLICANT: MCCORD/DALE, FINLAY (MMN)
INVENTOR: BUSCH, ALFRED (NNN)
TITLE OF INVENTION: DYE TRANSFER INHIBITING COMPOSITIONS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: THE PROCTER & GAMBLE COMPANY
STREET: 5299 SPRING GROVE AVENUE
CITY: CINCINNATI
STATE: OHIO
COUNTRY: USA
ZIP: 45217
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,249
FILING DATE: 05-MAY-1994
CLASSIFICATION: 510
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91202882.6
FILING DATE: 06-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: ALLEN, GEORGE W.
REGISTRATION NUMBER: 26,143
REFERENCE/DOCKET NUMBER: CM-395
TELECOMMUNICATION INFORMATION:
TELEPHONE: 513-627-5946
TELEFAX: 513-627-8118
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1060 base pairs
TYPE: nucleic acid
STRAND/ID: unknown
TOPOLOGY: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 10..924
US-08-232-249-1

Query Match 37.3%; Score 437.4; DB 1; Length 1060;
Best Local Similarity 70.4%; Pred. No. 1.3e-78;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CAGAGATGCGCTCTATCTCCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGAG 114
DB 5 CCAAGATGCGCTCTATCTCCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGAG 64
QY 115 CTTCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 174
DB 65 CTTCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 118
QY 175 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 234
DB 119 GCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 178
QY 235 AGCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 291
DB 179 AGCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 238
QY 292 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 351
DB 239 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 298
QY 352 GCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 411
DB 299 CTTCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 358
QY 412 CCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 471
DB 359 CCGCTTCCCGCGCGCAATGCGCTGCAATGCGCTGCAACCCCTGCGCTGCACTTCTCTGAG 418

QY 772 GCAAGG-----TGGACCGGAGCGCCACCTGACCTGCGCTGGT 813
Db 719 GCAACAGCTCTCCGGTCAACAGCTTACCAAGCAAGCAAGCTCCACCTCAACCT 778
QY 814 CGAGCCAGAC---GTCTCCCGGCGGCGAGTGGCTGACGCTTCAAGAGTGGCTCAGT 870
Db 779 CGAGCCCGGCGAGCTTCAAGCTTCCAGCGGCTGACGCTGAGAGGTGGCTCAGT 838
QY 871 GCGGTGGATGCGCTTCAAGCGGATCAACCACTGTGTCTTGGGACCACTGCGAAGT 930
Db 839 GCGGCGGATGGCTGGAGCGGCTGCAACCACTGCGTGGGAGCACTTGGAGAGAG 898
QY 931 TGAAGACTACTGCTGCGAGTGGCTTAAAC 961
Db 899 TTAATGACTGGTACATCACTGCTGTAGAC 929

RESULT 9

US-08-833-642A-1

; Sequence 1, Application US/08833642A

; Patent No. 5883066

; GENERAL INFORMATION:

; APPLICANT: Ivan M. A. J. Herbots et al.

; TITLE OF INVENTION: Liquid Detergent Compositions

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Jackie Ann Zurcher

; ADDRESS: Danmore & Shohl LLP

; STREET: 255 E. Fifth Street

; CITY: Cincinnati

; STATE: Ohio

; COUNTRY: USA

; ZIP: 45202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/833,642A

; FILING DATE: April 8, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Zurcher, J. A.

; REGISTRATION NUMBER: P42,251

; REFERENCE/DOCKET NUMBER: CWS51C

; TELEPHONE: (513) 977-8377

; TELEFAX: (513) 977-8141

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1060 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Humicola insolens

; STRAIN: DSM 1800

; FEATURE:

; NAME/KEY: mat peptide

; LOCATION: 73..927

; FEATURE:

; NAME/KEY: sig peptide

; LOCATION: 10..72

; NAME/KEY: CDS

; LOCATION: 10..927

; US-08-833-642A-1

Query Match 37.3%; Score 437.4; DB 2; Length 1060;
Best Local Similarity 70.4%; Pred. No. 1.3e-76;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CCAAGATGGCTCTACTCCGTTCTTCCGACAACTCGGCGGCTGACTTCCCTGCTGCTG 114
Db 5 CCAAGATGGCTTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 64
QY 115 CTTCCCGGCGAGTGGAGTGGCACTGCAAGATCTGGAATCTGCTGCAAGCGCTGCT 174
Db 65 CTTCCCGGCGAGTGGAGTGGCACTGCAAGATCTGGAATCTGCTGCAAGCGCTGCT 118
QY 175 GCGCTGGCGCGGAGGAGGCGCGCTGAGCAACCGGTCTAAGGTGGCACTTCC 234
Db 119 GCGCTGGCGCGGAGGAGGCGCGCTGAGCAACCGGTCTAAGGTGGCACTTCC 178
QY 235 AGGCTGCTCCGACTTCAATGTCAGTGGGCTGCA---CGGCGCTGCTGCTGCT 291
Db 179 AGGCTGCTCCGACTTCAATGTCAGTGGGCTGCA---CGGCGCTGCTGCTGCT 238
QY 292 GCGCTGCTCCGACTTCCCTGCGGCTGCAAGCAATCTGCTGCAAGCTTCCGCGCA 351
Db 239 GCGCTGCTCCGACTTCCCTGCGGCTGCAAGCAATCTGCTGCAAGCTTCCGCGCA 298
QY 352 GCAATGCGGCGGAGTGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
Db 239 CTAATGCGGCGGAGTGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
QY 412 CCGCTGCTCCGCTGCGGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
Db 359 CCGCTGCTCCGCTGCTGCGGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418
QY 472 GAAGTACCACTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
Db 419 GCAAGCAACCACTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478
QY 532 GCTGCAAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
Db 479 CTTCCCGGCTGCGGAGTGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
QY 592 GCGATTCCTTCCCGGCGGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
Db 539 GCGATTCCTTCCCGGCGGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
QY 652 AGCGCGCAACCGCACTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 711
Db 599 AGCGCGCAACCGCACTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 658
QY 712 GCTCCGCTGCAAGCTGCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 771
Db 659 GCAACGAGTGGCGGAGTGGCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 718
QY 772 GCAAGG-----TGGACCGGAGCGCCACCTGACCTGCGCTGGT 813
Db 719 GCAACAGCTCTCCGGTCAACAGCTTACCAAGCAAGCAAGCTCCACCTCAACCT 778
QY 814 CGAGCCAGAC---GTCTCCCGGCGGCGAGTGGCTGACGCTTCAAGAGTGGCTCAGT 870
Db 779 CGAGCCCGGCGAGCTTCAAGCTTCCAGCGGCTGACGCTGAGAGGTGGCTCAGT 838
QY 871 GCGGTGGATGCGCTTCAAGCGGATCAACCACTGTGTCTTGGGACCACTGCGAAGT 930
Db 839 GCGGCGGATGGCTGGAGCGGCTGCAACCACTGCGTGGGAGCACTTGGAGAGAG 898
QY 931 TGAAGACTACTGCTGCGAGTGGCTTAAAC 961
Db 899 TTAATGACTGGTACATCACTGCTGTAGAC 929

RESULT 10

US-08-140-008A-3

; Sequence 3, Application US/08140008A

; Patent No. 5914306

GENERAL INFORMATION:
APPLICANT: SVENDSEN, Allan
APPLICANT: VON DER OSTEN, Claus
APPLICANT: CLAUSEN, Ib Groth
APPLICANT: PATKAR, Shankar Arant
APPLICANT: BORCH, Kim
TITLE OF INVENTION: STABILIZED ENZYMES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: No. 59143060 No. 5914306disk of No. 5914306ch America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/140,008A
FILING DATE: 22-OCT-1993
CLASSIFICATION: 252
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3601,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 10..924
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 10..72
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 73..924
US-08-140-008A-3

Query Match 37.3%; Score 437.4; DB 2; Length 1060;
Best Local Similarity 70.4%; Pred. No. 1.3e-78;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CCAGATGCGCTTACTCCGTTCTTCGCAACCTTGACCGCTGCACTTCTCTGTCG 114
DB 5 CCAAGATGCGTCTCCCTCCCTCTCCGCGCTGAGCGCGCTGAGCGCGCTGAGCG 64
QY 115 CCGCGCGCGCGAGTGGGCGAGTCCGACGACGAGTCTACGCGTGCAGTGCAGCTTCG 174
DB 65 CCGTTCGCGCTGATGAGG-----GTCCACCGCTCTGAGAGCTGCGAGGCTTCTG 118
QY 175 GCGCTTGGCCCGGAGAGCGCGCTGAGCGACGAGTCTACGCGTGCAGTGCAGCTTCG 234
DB 119 GCGGCTGGGCGCAAGAGGCTCCGCTGACGAGCTCTTCTTCTGCAAGCGCACTTCG 178
QY 235 AGGCGCTGTGCGACTTCAATGTCAGTCTGAGCTGCAAA---CGGCGGCTGCGCTTCTCT 291
DB 179 AGCGTATCAAGGACTTCAAGTCAAGTCTGAGCTGAGCGCGCGGCTGCTGCGCTACGCT 238
QY 292 GCGCGAGCGAGTCCCGTGGGCGGTGAGAGCATCTGCGCTTACGCGTTCGCGCGAGCA 351
DB 239 GCGCGAGCGAGTCCCGTGGGCGGTGAGAGCATCTGCGCTTACGCGTTCGCGCGAGCA 298

QY 352 GCAATGCGCGCGGCTCCGAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
DB 299 CTAATTCGCGGAGGATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
QY 412 CCGGTCGCGCGCGGAGAGCAATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 471
DB 359 CCGGTCGCTGCTGCTGAGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418
QY 472 GAATGACAGTGTGATATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 531
DB 419 GAGAGACCACTTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 478
QY 532 GCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
DB 479 CTCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
QY 592 GCGATTCCTTCCCGCGCGCTCAAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
DB 539 GCGATTCCTTCCCGCGCGCTCAAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
QY 652 AGCGTGAACCCGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 711
DB 599 AGCGTGAACCCGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 658
QY 712 GCTGCGCGGAGAGGAG 771
DB 659 GCAAGGAG 718
QY 772 GCAAGGAG 813
DB 719 GCAAGGAG 778
QY 814 CCGGCGAGAG---GCTCCCGCGCGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 870
DB 779 CAGGCGCGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 838
QY 871 GCGGAG 930
DB 839 GCGGAG 898
QY 931 TGAAGACTTACTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 961
DB 899 TGAAGACTTACTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 929

RESULT 11
US-08-389-423-1
Sequence 1, Application US/08389423
Patent No. 5948672
GENERAL INFORMATION:
APPLICANT: Rasmussen, Grethe
APPLICANT: Mikkelson, Jan Moller
APPLICANT: Schuelein, Martin
APPLICANT: Patkar, Shankar A.
APPLICANT: Hagen, Fred
TITLE OF INVENTION: A Cellulase Preparation Comprising an
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: No. 59486720 No. 5948672disk of No. 5948672th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/389,423
FILING DATE: 14-FEB-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3469, 214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Humicola insolens
STRAIN: DSM 1800
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 73..924
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 10..72
FEATURE:
NAME/KEY: CDS
LOCATION: 10..924
US-08-389-423-1

Query Match 37.3%; Score 437.4; DB 2; Length 1060;
Best Local Similarity 70.4%; Pred. No. 1.3e-78;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

55 CCAGAGATGCGCTTACTCCCGTTCTGCGCAACCCCTGCGCTGCACTTCTCTGCTGCG 114
5 CCAGATGCGCTTCTCCCGCTTCTCCCGTCCCGTCCCGTCCCGTCCCGTCCCGTCCCGT 64
115 CCTCCGCGCGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 174
65 CCTCCGCGCGCTGATGCGAG-----GTCCACCCGCTACTGAGACCTGCTGCAAGCCCTTGT 118
175 GCGCTGCGCGCGAGAGCGCGCGCTGCGCAACCGGCTGCAAGCGGCTGCAAGCGCACTTCC 234
119 GCGCTGCGCGCGAGAGAGCGCTCCCGTGAACCGCTGCTCTTCTTCCGCAAGCGCACTTCC 178
235 AGCGCTGCTCGCACTTCAATGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 291
179 AGCGTATGCGAGCTTCAAGCGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 238
292 GCGCGACGAGACTCCCTGCGCGTGAACGACGACATGCTGCTGCTGCTGCTGCTGCTGCTG 351
239 GCGCGACGAGACCCCATGAGCTGGAACGACGACATGCTGCTGCTGCTGCTGCTGCTGCTG 298
352 GCATGCGCGCGCGGATCGCAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411
299 CATATGCGCGCGAGATGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
412 CCGTCCCGTGCAGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471
359 CCGTCCCGTGCAGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418
472 GAAGTACGAGTTCGATGCGCAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 531
419 GCGAGACGAGCTTCAATGCTGCAATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 478
532 GCTGCGAGTTCGCGCGCGCTCCCGCGCGCTCAATAGCGCGCGCGCGCGCGCGCGCGCGCGCG 591
479 CTCCCGAGTTTCGCGCGCTGCG 538
592 GCGATCTCTTCCCGCGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 651
539 GCGATCGATTCCCGCGCGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598

652 ACGCGCAACCCGACGTTTCCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 711
599 ACGCGCAACCCGACGTTTCCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 658
712 GCTCCGCGTCAAGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771
659 GCGCGAGTTCG 718
772 GCAACGCG-----TGGCACCGGAGCGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 813
719 GCAACGCTCTCCGTTCAACGCTTACGAGCAACGAGCAACGAGTGCAGTGCAGTGCAGTGCAG 778
814 CCGCGCGAGC-----GTCTCCG 870
779 CGAGCCGCGAGTTCAGGCTTACGAGCTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 838
871 GCGGTGCGATCGGCTTCAAGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930
839 GCGCGCGAGTTCGAGCG 898
931 TGAACGACTACTTCCGAGTGCCTTAAAC 961
899 TTATGATGTAACCATCAGTGCCTGTAGAC 929

RESULT 12
US-08-816-915-7
Sequence 7, Application US/08816915
Patent No. 6060305
GENERAL INFORMATION:
APPLICANT: Royer, John C
APPLICANT: Moyer, Donna L
APPLICANT: Yoder, Wendy T
APPLICANT: Shuster, Jeffrey R
TITLE OF INVENTION: NON-TOXIC, NON-TOXIGENIC, NON-PATHOGENIC
TITLE OF INVENTION: FUSARIUM EXPRESSION SYSTEM
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6060305 No. 6060305disk of No. 6060305th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,915
FILING DATE: 13-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agtis Dr. Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4216, 240-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..924
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 73..924

```

;      FEATURE:      sig_peptide
;      NAME/KEY:      ;
;      LOCATION:      10..72
US-08-816-915-7

```

Query Match	37.3%;	Score 437.4;	DB 3;	Length 1060;
Best Local Similarity	70.4%;	Pred. No. 1.3e-78;		
Matches 655;	Conservative 0;	Mismatches 246;	Indels 30;	Gaps 4

QY	55	CCAGGATGGGCTCTACTCCGCTCTTTCGACAAACCTGGAGCGGCTGCACTTCCTGTGAG	11
Dp	5	CCAGATAGGCTTCTCCCTCCTCCTCCGCTCCGCTGTTTGAGCGCCTCTGCGGTATTTG	64
QY	115	CTCTCCGCGCACTGAGTGGCACTGCAAGAGTACTGAGACTGTCTGACACCGCTGT	174
Dp	65	CCCTTTCGCGTATGGACG-----GTCCACCGCTATCTGGGACCTGTGACACCTTGT	118
QY	175	GCGCTTGTGGCTCCGGGAAAGCGGCTGTACGACACCGGTCTTACGCGTGGAGTCCACTTCC	234
Dp	119	GCGGCTGGGCAAGAAAGGCTCTCCGTAAACAGCGCTGTCTTTTCTTGCAAGCCCACTTCC	178
QY	235	AGCGCTGTCCGACTTCAATTAATGTCACTGGGGCTGCA-----CGGCGGCTCCGGCTACTCT	291
Dp	175	AGCGTATCAACGACTTTCAGCGCAAGCTCCGGCTCCGAGCGCGGCGGAGTCACTTCTGT	238
QY	292	GCGCGCAACGACTCCTCTGGGCGGTGAACGACATCTCCCTACCGGCTCTGCGCGGACGA	351
Dp	239	GCGCGCAACGACCCCAATGGGCTGTGAACGACATTTCCGCGTGGTTTGTGCGACCT	298
QY	352	GCAATGCTCCGCGGCTTCGAAATCTCTGTGTGTCTGCGCTGTACGCGCTCACTTCACTT	411
Dp	299	CTATTGTCGCGCAACATGAGAGCGGAGTGTCTGTCTGCGCTGTCAAGACTCACTTCACT	358
QY	412	CCGGTCCCGTCCGCGCGCAACAGATATGTGTGTGAGCTTCAACGACACCTGGCGGACCTGG	471
Dp	359	CCGGTCCCTGTCTGTGGCAAGATATGTGTGTGTCGACTCAACCGACACCTGGGGTATCTTG	418
QY	472	GAAGTAAACGATTTGATATGTGACATGCCCCGCGCGGCGGTGGAGACTTTTCAACGGCTCA	531
Dp	419	GAGCAACCACTTGTATCTCAATCTCCCGCGCGGCGGTGGAGACTTTTGAAGGATGCA	478
QY	532	GCTGCGAGTTGGGCGGCTCTCCCGGCGCTCAATACGCGGCAATTTGTGTGCGACGAGT	591
Dp	479	CTCCCAAGTTCCGCGGCTGTCTCCCGCGACGCGCTAACGCGGCACTCTGTCCCGACGAGT	538
QY	592	GCGATTCCTTCCCGCGCGCGCTCAAGCCCGCGCTCCGACGTGGGTGTGATGTTTCAAGA	651
Dp	539	GCGATGTGTTCCCGAGCGCCTTAAGCCCGGCTGTCACTACGGCGCTTCACTGTATCAAGA	598
QY	652	ACGCGCAACACCCGACTTACGCTTCCAGCAGGTGACAGTCCCGCGGAGATGTTGGCC	711
Dp	599	ACGCGCAACATCCAGGTTTCAGCTTCCGTCAAGTTCAGTCCCGCGGAGCTGTGCTCTC	658
QY	712	GCTCCGCTGTCAAGCGCAACGACACTCAAGCTTCCCGCTTTCACACCCCGCAAGCGGTG	771
Dp	659	GCAACCGGATGCGCGCGCAACGACGAGCGCAACTTCCCTGCGTCAATATCCCTCAAGCA	718
QY	772	GCAACGG-----TGGCAACCGGAGCGCCACGCTGCACTGCGCTGGGT	813
Dp	719	GCAACAGCTTCTCCGCTTAACAGGCTTACAGAGCAACGACACACTTCACTTCAACACT	778
QY	814	CGGCGCAAGAC---GTCTCCCGCGCGCGGCAAGTGGTGCAGGCTCAAGATGGGCTAGT	870
Dp	779	CGAGCCCGCAAGTCAAGCTTCAAGCTCCGACCGGCTGCACTGCTGAGAGGTGGGCTAGT	838
QY	871	GCGGTGGCATCGGCTTTCAGCGGATGCAACACTGTGTCTTGTGACCACTTCCAGAAAT	930
Dp	839	GCGGCGGCAATGGCTGGAGGGCTGCAACACTGTGCTGTGGAGCACTTTCAGAAAG	898
QY	931	TGAACGACACTACTGCGAGTGGCTCTTAAAC	961
Dp	899	TTAATGATCGGATCAATCAAGTGTCTGTAGAC	929

RESULT 13
US-09-189-060B-55
; Sequence 55, Application US/09189060B
Patent No. 6720000

```
; NAME/KEY: sig_peptide
; LOCATION: (10)...(72)
US-09-189-060B-55
```

Query Match	37.3%;	Score 437.4;	DB 3;	Length 1060;
Best Local Similarity	70.4%;	Pred. No. 1,3e-78;		
Matches	655;	Conservative	0;	Mismatches 246; Indels 30; Gaps 4

Qy	55	CCAGAGATGCGCTCATCTCCCGTTCTTTGGCAACACCTGGCCGCTGCACCTTCTCTGGTGG	114
Db	5	CCAGATATGGTTCTCCCGCCCTCTCCCGCCCGCCCTTGGTGGCCGCTTGGCCGCTTGG	64
Qy	115	CCCTCCCGCGCAGATGGCGACATGGCCAGTCCACAGAGATACTGGGACCTGGTCCAGACCGTGGT	174
Db	65	CCCTTTCGCGCTGATATGGAG-----GTCCACCGGCTACTGGGATCTGTGCAGACCTTGGT	118
Qy	175	GCGCTTGGCCCGGAGAGGCCCGCGCTGCAGCCCAACCGGTCCTACGCGTGCATATGCCAATTCC	234
Db	119	GCGCGTGGCGCCAGAGAGCGCTCCCGTGAACACAGCTGTCTTTCTTCCGACAGCCAACTTCC	178
Qy	235	AGGCGCTGTGCGCATTCATATGTCAGATGCGGCGCTGCAG---CGGCGGCGTGGCGCTATCTCT	291
Db	179	AGGCTATCACGAGATTTCACGACCAAGTCCGCGCTGCGAGCCGGGGGTGTGCGCTACTGGT	228
Qy	292	GCGCGGACCGACATCTCCCTGGGCGGTGAACGACATATCTGCGCTACCGGCTTTCGCGGACGA	351
Db	239	GCGCCGACCAACACCCCATATGGGCTGTGAACGACATTCGCGCTGGTTTGTGCGACCT	298
Qy	352	GCAATGCGCGCGGCGGCTCGAAATCTCTGCGGTGCTGGCGGCTGTACAGGCTCTTCACTT	411
Db	299	CTATTGGCGGACGAAATGAGGGGGGCTGGGTGCTGGCGCTGTACAGCTCACTTCACTTCACT	358
Qy	412	CCGATCTCCGTCGCGCGACAGACATATGGTGTGACGTCAACAGACATGCGCGGACACTGG	471
Db	359	CCGGTCTGTTTGGCGGACAGAAATGGTGGTGCATTCACACAGACATGCGCGGTATCTTG	418
Qy	472	GAAATGACCACTTTCATATTCGCAATGCGCGGCGGCGGCGGCGGCGGCAATCTTCAACGCTGCA	531
Db	419	GAGAGACCACTTTCATATTCGCAATCTCCCGCGGCGGCGGCGGCGGCGGCAATCTTCAACGAGTGA	478
Qy	532	GCTCCGATTTGCGCGGCGCTCCCGCGGCGCTCAATAAGCGGCAATTTGTGTGCGCGACAGT	591
Db	479	CTGCCCAATTTGGGCGGCTGTGGCCGCGACAGGCTTACAGGGGCGCATCTGTGTCCGCAACGAGT	538
Qy	592	GCGATTCCTTCCCGCGCGCTCAAGCTCGGCGTGCAGATGGCGGTTTGAATCTGTTCAGAC	651

Db 539 GCGATCGTTCCCGACCGCCTCAAGCCGCGCTGCTACTGGGCTTTCAGACTGGTTCAAGA 598
Qy 652 AGCCGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
Db 599 AGCCGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
Qy 722 GCTCGGCTGCAAG 771
Db 659 GACCGGATGCGCGCAACAG 718
Qy 772 GCAACGG-----TGACACGAG 813
Db 719 GACCGAGCTTCGCGTCAACAG 778
Qy 814 GCGGACAGAC--GTCTCCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
Db 779 GAGCGCGGAGCTCAGACTCCAG 838
Qy 871 GCGGAGAGATCGGCTTCAAG 930
Db 839 GCGGAGAGATGAGCTGAG 898
Qy 931 TGAAGACTACTACTCGAG 961
Db 899 TTAATGACTGTATCAATCAAG 929

RESULT 14

US-09-230-665-1
Sequence 1, Application US/09230665
Patent No. 6322595
GENERAL INFORMATION:
APPLICANT: Boyer, Stanton L
TITLE OF INVENTION: Detergent Composition Comprising Two Cellulase
FILE OF INVENTION: Components, with and without a Cellulose-Binding Domain
PATENT REFERENCE: 6191 Sequence Listing (8 Sequences)
Patent No. 6322595
CURRENT APPLICATION NUMBER: US/09/230,665
CURRENT FILING DATE: 1999-01-29
EARLIER APPLICATION NUMBER: 60/023,125
EARLIER FILING DATE: 1996-07-30
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1060
TYPE: DNA
ORGANISM: Humicola insolens
FEATURE:
NAME/KEY: CDS
LOCATION: (10)..(924)
FEATURE:
NAME/KEY: mac_peptide
LOCATION: (73)..(924)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: (10)..(72)
US-09-230-665-1

Query Match 37.3%; Score 437.4; DB 4; Length 1060;
Best Local Similarity 70.4%; Pred. No. 1.3e-78;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

Qy 55 CGAGATGCGCTTACTCCCGCTTCTGCAACACCTGCGCGCTGAGACTTCTCTGGTGG 114
Db 5 CGAGATGCGCTTCTCCCGCTTCTGCAACACCTGCGCGCTGAGACTTCTCTGGTGG 64
Qy 115 CCTCGCGGCGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 174
Db 65 CCTCGCGGCGTATGCGAG-----GTCAACCGGCTACTGGAATGCTGCAAGCCTTGT 118
Qy 175 GCGCTGGCGCGGAGAGCGCGCTGAGCGCAACCGCTTAAGCTGCGAGTGCAGACTTCC 234
Db 119 GCGCTGGCGCGGAGAGCGCGCTGAGCGCAACCGCTTCTTCTGCAAGCGCACTTCC 178

Qy 235 AGGCTGTCCGACTTCAATGTCAGTGGGCTGCA--CGAGCGCTCGGCTTACTCT 291
Db 179 AGGCTATCAAGCACTTCAAGCGCAAGTCCGGCTGCGAGCGGCGAGTCTGCTCACTCGT 238
Qy 292 GCGCGACCAATCTCCCTGGGCGGTGAACAGCAATCTGCTTAAGGCTTGGCGCGAGCA 351
Db 239 GCGCGACCAATCTCCCTGGGCGGTGAACAGCAATCTGCTTAAGGCTTGGCGCGAGCA 298
Qy 352 GCATCGCGCGGAGTCCGATCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
Db 299 CATTCGCGGAGCAAGTGAAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
Qy 412 CGGTCGCGTGGCGCGGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
Db 359 CGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418
Qy 472 GAAGTACAGTGTGATATGCGATGCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 531
Db 419 GCAAGCACTACTGATGATCTCAATCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 478
Qy 532 GCTCGAGTTCGCGGCGCTCCCGCGCTCAATACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 591
Db 479 CTCCCAAGTTCGCGGCGCTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 538
Qy 592 GCGATTCCTTCGCGCGCGGCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 651
Db 539 GCGATGCTTCGCGCGCGGCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 598
Qy 652 AGCGGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
Db 599 AGCGGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
Qy 712 GCTCGGCTGCAAG 771
Db 659 GCAACCGATGCGCGCGGAG 718
Qy 772 GCAACGG-----TGACACGAG 813
Db 719 GACCGAGCTTCGCGTCAACAG 778
Qy 814 GCGGACAGAC--GTCTCCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
Db 779 GAGCGCGGAGCTCAGACTCCAG 838
Qy 871 GCGGAGAGATCGGCTTCAAG 930
Db 839 GCGGAGAGATGAGCTGAG 898
Qy 931 TGAAGACTACTACTCGAG 961
Db 899 TTAATGACTGTATCAATCAAG 929

RESULT 15

US-09-189-028-1
Sequence 1, Application US/09189028
Patent No. 6423524
GENERAL INFORMATION:
APPLICANT: Raemussen, Gretle
APPLICANT: Mikelseth, Jan Moller
APPLICANT: Schuelein, Martin
APPLICANT: Parker, Shankant A.
APPLICANT: Hagen, Fred
TITLE OF INVENTION: A Cellulase Preparation Comprising an
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: No. 64235240 No. 6423524disk of No. 6423524th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America

ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/189,028
APPLICATION NUMBER: US/09/189,028
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/389,423
FILING DATE: 14-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3469.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: cDNA
TOPOLOGY: linear
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: *Humicola insolens*
STRAIN: DSM 1800
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 73..924
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 10..72
FEATURE:
NAME/KEY: CDS
LOCATION: 10..924
US-09-189-028-1

Query Match 37.3%; Score 437.4; DB 4; Length 1060;
Best Local Similarity 70.4%; Pred. No. 1.3e-78;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CAGAGTGGCTTACTCCCGTCTTCGACAACTGCGCGCTGCACTTCCTGTGTG 114
DB 5 CCAAGATGGCTTCCTCCCTCTCCGTCGCGCTGTGTGTGCGCCCTGCGGTGTG 64

QY 115 CTTCCCGGCGGAGTGGCGAGTGGCCAGTCCAGATCTGGAATCTGTGCAAGCGTCT 174
DB 65 CCTTTCGCGCTGATGGGAG-----GTCCACCGCTTCTGGGATGTGTCMAAGCTTGT 118

QY 175 GCGCTTGGCCGAGAAAGCGCGCTGAGCAACCGATCTACGGGTGGCATGCCAATCTCC 234
DB 119 GCGCTTGGCCGAGAAAGCGCTTCCGTAACCAAGCTGTCTTTCTGCAACGCACTTCC 178

QY 235 AGCGCTGTCCGACTTCAATGTCCAGTGGGCTGCAA---CGCGGCTCGGCTTACTCT 291
DB 179 AGCGTATCAAGGACTTGAAGCCAGTCCGAGCTGCGAGCGCGGCTGTGCGCTACTCGT 238

QY 292 GCGCCGACCAACTCTCTCGGCGGTGAACGACATCTGCTACGCTTGGCGCGACGA 351
DB 239 GCGCCGACCAAGCCCAATGGGCTGTGAACGACATCTGCGGCTCGGTTTGTGCGCACT 298

QY 352 GCATGCGCGGCGGAGTCTGAGTGTGCTGCGCTGCTACAGGCTCACTTCACTT 411
DB 299 CTATTGCGGCGAGCAATGAGCGGCTGTGTGCTGCGCTGCTACAGCTCACTTCACT 358

QY 412 CCGGTCCTGTGCGGCAAGCAATGTGTGCTGCTACAGCACTGCGCGGCTGATCTTG 471
DB 359 CCGGTCCTGTGCTGCGCAAGCAATGTGTGCTGCTACAGCACTGCGCGGCTGATCTTG 418

QY 472 GAATTAACCACTTCATATCGCATGCCGCGCGCGGCTGAGATCTTCAACGCTGCA 531
DB 419 GAGCAACCACTTCATATCGCATGCCGCGCGCGGCTGAGATCTTCAACGCTGCA 478

QY 532 GCTCCAGTTCCGCGGCTCCCGGCGCTCAATACGCGGCACTTGTGCGCGGAGCACT 591
DB 479 CTCCCAATTCGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 538

QY 592 GCGATTCCTTCCCGGCGGCTCAAGCCCGGCTGCGAGGCGGCTTGAATGTTCCAGA 651
DB 539 GCGATTCCTTCCCGGCGGCTCAAGCCCGGCTGCGAGGCGGCTTGAATGTTCCAGA 598

QY 652 AGCGCAACCTTACGTTTCAAGTTCAGAGGTCAGATGTCGCGGCGGAGATGTTGCC 711
DB 599 AGCGCAACCTTACGTTTCAAGTTCAGAGGTCAGATGTCGCGGCGGAGATGTTGCC 658

QY 712 GCTCCGCTGAGAGCGGAGGAGCTTCAGCTTCCCGGCTTCAACCCCGGAGGCGG 771
DB 659 GCGCGGATTCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 718

QY 772 GCACGCG-----TGGACCGGAGAGCGGCAAGTGAATGCGGCTGAGGT 813
DB 719 GCACGAGCTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 778

QY 814 CGGCGCAAGC---GTCTCCGCGGCGGCGGCTGAGGCTGCAAGTGTGAGTGTGAGT 870
DB 779 CGAGCGCGGAGTTCAGGCTTCAAGCTCCAGCGGCTGCACTGTGAGAGGTGTGAGT 838

QY 871 GCGGTGGATGGCTTCAAGGAGTGAACCACTGTGTCTGTGAGCACTGCGAGAGT 930
DB 839 GCGGTGGATGGCTTCAAGGAGTGAACCACTGTGTCTGTGAGCACTGCGAGAGT 898

QY 931 TGAACGATCTACTGCGAGTCCCTTAAC 961
DB 899 TTAATGATGTGATCATGATGCTGTAGAC 929

Search completed: July 7, 2004, 11:41:08
Job time: 123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 10:19:37 ; Search time 4330 Seconds
(without alignments)
1307.995 Million cell updates/sec

Title: US-10-007-521-11

Sequence: 1 GAGCAGCAGCCTCCAGCTG.....AAAAAAAAAAAAAAAAAAAA 1174

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1174	100.0	1174	US-10-007-521-11	Sequence 11, Appl
2	453.6	38.6	1261	US-09-261-829-23	Sequence 23, Appl
3	437.4	37.3	1060	US-09-735-787-1	Sequence 1, Appl
4	437.4	37.3	1060	US-10-138-870-1	Sequence 1, Appl
5	377.2	32.1	890	US-10-369-893-27891	Sequence 27891, A
6	377.2	32.1	1257	US-10-416-828-7	Sequence 7, Appl
7	377.2	32.1	1257	US-10-432-290-18	Sequence 18, Appl
8	369.8	31.5	1154	US-10-007-521-7	Sequence 7, Appl
9	353.6	30.1	936	US-08-841-636A-30	Sequence 30, Appl
10	337	28.7	894	US-10-007-521-3	Sequence 3, Appl
11	334.8	28.5	927	US-10-007-521-5	Sequence 3, Appl
12	315.2	26.8	1473	US-09-735-787-3	Sequence 3, Appl
13	315.2	26.8	1473	US-10-138-870-3	Sequence 3, Appl
14	313.6	26.7	885	US-10-007-521-23	Sequence 23, Appl

15	302	25.7	1132	US-10-007-521-21	Sequence 21, Appl
16	289	24.6	672	US-10-432-290-39	Sequence 39, Appl
17	288.2	24.5	1423	US-10-007-521-9	Sequence 9, Appl
18	274.4	23.4	960	US-10-007-521-1	Sequence 1, Appl
19	260.2	22.2	1043	US-10-432-290-13	Sequence 13, Appl
20	254.6	21.7	913	US-10-007-521-13	Sequence 13, Appl
21	228	19.4	1041	US-10-432-290-12	Sequence 12, Appl
22	213.8	18.2	425	US-10-007-521-25	Sequence 25, Appl
23	210.8	18.0	1031	US-10-007-521-19	Sequence 19, Appl
24	210.8	18.0	1048	US-10-007-521-17	Sequence 17, Appl
25	184	15.7	1017	US-10-432-290-8	Sequence 8, Appl
26	181	15.4	1083	US-10-432-290-10	Sequence 10, Appl
27	158.8	13.5	1101	US-10-432-290-6	Sequence 6, Appl
28	153.6	13.1	1017	US-10-432-290-2	Sequence 4, Appl
29	150.2	12.8	808	US-10-007-521-15	Sequence 15, Appl
30	150.2	12.8	808	US-10-007-521-37	Sequence 37, Appl
31	109.8	9.4	1177	US-10-007-521-39	Sequence 39, Appl
32	98.6	8.4	153	US-10-007-521-31	Sequence 31, Appl
33	91	7.8	180	US-10-007-521-41	Sequence 41, Appl
34	85.2	7.3	171	US-10-007-521-51	Sequence 51, Appl
35	84.2	7.2	225	US-10-007-521-31	Sequence 36, Appl
36	83.2	7.1	887	US-08-841-636A-36	Sequence 71, Appl
37	80.8	6.9	165	US-10-007-521-77	Sequence 71, Appl
38	70.8	6.0	159	US-10-007-521-71	Sequence 57, Appl
39	69.8	5.9	147	US-10-007-521-57	Sequence 57, Appl
40	69.2	5.9	177	US-10-007-521-33	Sequence 33, Appl
41	66.2	5.6	413	US-10-437-963-86254	Sequence 86254, A
42	66.2	5.6	608	US-10-424-599-75014	Sequence 75014, A
43	66	5.6	419	US-10-437-963-20567	Sequence 20567, A
44	66	5.6	682	US-09-764-847-20	Sequence 20, Appl
45	66	5.6	682	US-10-092-154-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-10-007-521-11
; Sequence 11, Application US/10007521
; Publication No. US20030054539A1
; GENERAL INFORMATION:
APPLICANT: Schleen, Martin
Lassen, Soren F.
Kauppinen, Markus S.
Lange, Lene
Nielsen, Rudy I.
Ihara, Michiko
Takagi, Shinobu
TITLE OF INVENTION: No. US20030054539A1e1 Endoglucanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSER: No. US20030054539A10 No. US20030054539A1disk of No. US20030054539A1
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366-200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..956
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-007-521-11

Query Match 100.0%; Score 1174; DB 15; Length 1174;
Best Local Similarity 100.0%; Pred. No. 1e-265;
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGACCCCTCAAGCTGACAGTTTCAACCCCGCTCTCTTTCTTCCGCCCAAGA 60
DB 1 GAGCAGACCCCTCAAGCTGACAGTTTCAACCCCGCTCTCTTTCTTCCGCCCAAGA 60
QY 61 TGGGCTCTACTCCCGTTCTTCCCAACCTTGGCCCTGCACTTCTCTGATGCTCCG 120
DB 61 TGGGCTCTACTCCCGTTCTTCCCAACCTTGGCCCTGCACTTCTCTGATGCTCCG 120
QY 121 CGGCGAGTGGCGAGTGGCGAGTGCAGATATCTGGGAGCTGCGAAGCGGTGCTT 180
DB 121 CGGCGAGTGGCGAGTGGCGAGTGCAGATATCTGGGAGCTGCGAAGCGGTGCTT 180
QY 181 GGCCTCGAGAGCGCGCTGACGCAACCGGTCTAGCGGTGATGCAATCTTCAGCGCC 240
DB 181 GGCCTCGAGAGCGCGCTGACGCAACCGGTCTAGCGGTGATGCAATCTTCAGCGCC 240
QY 241 TGTCCGACTCAATGTCAGTGGCGGTGGCAAGCGGCGCTGGGCTTACTCTGCGGACC 300
DB 241 TGTCCGACTCAATGTCAGTGGCGGTGGCAAGCGGCGCTGGGCTTACTCTGCGGACC 300
QY 301 AGACTCCCTGGGCGGTGAAACCAATCTGCTCAACCGCTTGGCGGAGCAATCGCGG 360
DB 301 AGACTCCCTGGGCGGTGAAACCAATCTGCTCAACCGCTTGGCGGAGCAATCGCGG 360
QY 361 GCGGATCGAATCTTGT 420
DB 361 GCGGATCGAATCTTGT 420
QY 421 TGGCGGAGAGCAATGTCAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATG 480
DB 421 TGGCGGAGAGCAATGTCAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATG 480
QY 481 AGTTGATATGTCAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 540
DB 481 AGTTGATATGTCAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 540
QY 541 TGGGCGCTTCCCGCGGTGAAACCAATCTGCTCAACCGCTTGGCGGAGCAATCGCGG 600
DB 541 TGGGCGCTTCCCGCGGTGAAACCAATCTGCTCAACCGCTTGGCGGAGCAATCGCGG 600
QY 601 TGGCGGCGCTTCCCGCGGTGAAACCAATCTGCTCAACCGCTTGGCGGAGCAATCGCGG 660
DB 601 TGGCGGCGCTTCCCGCGGTGAAACCAATCTGCTCAACCGCTTGGCGGAGCAATCGCGG 660
QY 661 ACCGAGCTTCAAGTTCAGAGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 720
DB 661 ACCGAGCTTCAAGTTCAGAGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 720
QY 721 GCAAGCGCAAGCAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 780
DB 721 GCAAGCGCAAGCAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 780

QY 781 GCAAGCGCAAGCAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 840
DB 781 GCAAGCGCAAGCAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 840
QY 841 GTGGCTGCAAGTTCAGAGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 900
DB 841 GTGGCTGCAAGTTCAGAGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 900
QY 901 CTTGTGTCTTGGCAACCTTGGCGGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 960
DB 901 CTTGTGTCTTGGCAACCTTGGCGGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 960
QY 961 CAGTTTTCAGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 1020
DB 961 CAGTTTTCAGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 1020
QY 1021 TTTTGAAGCGCTCAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 1080
DB 1021 TTTTGAAGCGCTCAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 1080
QY 1081 TCTCAGACCGACTTGGCGGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 1140
DB 1081 TCTCAGACCGACTTGGCGGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 1140
QY 1141 AA 1174
DB 1141 AA 1174

RESULT 2

US-09-261-329-23
Sequence 23, Application US/09261329
Publication No. US2003092097A1
GENERAL INFORMATION:
APPLICANT: Andersen, Kim
APPLICANT: Schuelin, Martin
APPLICANT: Christiansen, Lars
APPLICANT: Damgaard, Bo
APPLICANT: Von der Osten, Claus
TITLE OF INVENTION: Cellulase Variants
FILE REFERENCE: 4887.204-US
CURRENT APPLICATION NUMBER: US/09/261,329
CURRENT FILING DATE: 1999-03-03
EARLIER APPLICATION NUMBER: 1013/96
EARLIER FILING DATE: 1996-09-17
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 23
LENGTH: 1261
TYPE: DNA
ORGANISM: Humicola grisea

Query Match 38.6%; Score 453.6; DB 10; Length 1261;
Best Local Similarity 74.9%; Pred. No. 1.5e-96;
Matches 583; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

QY 26 TTTTCAACCCCGCTCTTTTCTTGGACCCCGAGAGTGGCTTACTCCCGTTCTTGAC 85
DB 18 TTTTATTCACCGCTCAATTATTTCAAACTTCATATGCGCTGCTCTTATTTTCGAC 77
QY 86 AACCTGGGCGGTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 145
DB 78 GGCCTGGGCGGTGCTTCCCTTGGCGACTGCGCGCATGCGA-----AGTGCAC 131
QY 146 GAGATCTGAGACTGTCAGAGCGTGTGTGCTTGGCCCTCGGAGAGCGCGCTGACCA 205
DB 132 CAGATCTGAGACTGTCAGAGCGTGTGTGCTTGGCCCTCGGAGAGCGCGCTGACCA 191
QY 206 ACGGCTGAGCGGTGAGTGCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 265
DB 192 GCGCTTCTCACTTGTGACACCGCAATTCAGAGGATTCAGAGCGCGCAATTCAGAGG 251

PATENT NO. US20010036910A1

SEQUENCE LISTING INFORMATION:

GENERAL INFORMATION:

APPLICANT: Rasmussen, Grethe
Mikkelsen, Jan Moller
Schuleid, Martin
Packar, Shankant A.
Hagen, Fred

TITLE OF INVENTION: A Cellulase Preparation Comprising an Endoglucanase Enzyme

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESS: No. US20010036910A1O No. US20010036910A1disk of No. US20010036910A1

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/735,787

FILING DATE: 13-Dec-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/189,028

FILING DATE: <Unknown>

AUTHOR/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 3469, 214-US

1	TELECOMMUNICATION INFORMATION:	
2	TELEPHONE: 212-867-0123	
3	TELEFAX: 212-878-9655	
4	INFORMATION FOR SEQ ID NO: 1:	
5	SEQUENCE CHARACTERISTICS:	
6	LENGTH: Humicola insolens	
7	TYPE: nucleic acid	
8	STRANDEDNESS: single	
9	TOPOLOGY: linear	
10	MOLECULE TYPE: cDNA	
11	HYPOTHEetical: NO	
12	ORIGINAL SOURCE:	
13	STRAIN: DSM 1800	
14	FEATURE:	
15	NAME/KEY: mat_peptide	
16	LOCATION: 73..924	
17	FEATURE:	
18	NAME/KEY: sig_peptide	
19	LOCATION: 10..72	
20	FEATURE:	
21	NAME/KEY: CDS	
22	LOCATION: 10..924	
23	SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
24	US-09-735-787-1	
25	Query Match	37.3%; Score 437.4; DB 9; Length 1060;
26	Best Local Similarity	70.4%; Pred. No. 9,2e-93;
27	Matches	653; Conservative 0; Mismatches 246; Indels 30; Gaps 4
28		
29	55	CGAGATGAGCTCACTCCGTTCTTGTCACAACTGACCTGACCTGACCTCTCTGATCG 114
30	5	CGAAGATGAGCTCTCCCCCTCCCTCCCGTCGCGCTTGATGAGCGCCCTGCGGATGAG 64
31		
32	115	CTTCGCGGCGCAGTGGCAGTGGCAGTCCAGATATCTGGAACTGCTGCAAGCCGTGCT 174
33	65	CCCTTCGCGCTGAGTGGAG-----GTCCACCCGCTACTGGAATGCTGCAAGGCTTCGT 118
34		
35	175	GCGCTGGCGCGGGAAGGCGCGCGTCAGCGCAACCGTGTACAGGTGCGCAATCGCACTTCC 234
36	119	GCGGCTGGCGCAAGAGCTCCCGTGAACAGCTGTCTTTTCTTGCAACGCAACTTCC 178
37		
38	235	AGGCGCTGTCGCACTTCAATGTCAGTCCGACTGCA---CGGCGCTGCGCACTTCT 291
39	179	AGGTATACAGGATCTTGACAGCAAGTCCGCTCGGACCGGCGGTGTGCTGCTACTCT 238
40		
41	292	GCGCGACGACGACTCTCCCTGGCGGTGAAGCAATCTGCGCTTACGCGCTTCCGCGACGA 351
42	239	GCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 298
43		
44	352	GCAATGCGCGGCGGCTCGCAATCTCTGAGTGCAGCGCGCTGCTACGCGCTACCTTACTT 411
45	299	CTATTGCGCGGACGATGAGGCGGCGCTGAGTGTGGCTGCTGCTGCTGCTGCTGCTGCTG 358
46		
47	412	CCGCTCCCGTGCACGCGGACGAATGATGATGATGATGATGATGATGATGATGATGATGAT 471
48	359	CCGCTCTGTTGCTGCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
49		
50	472	GAAATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531
51	419	GAGCAACCACTTGTGATTCACATCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 478
52		
53	532	GCTGCGAGTTCGCGCGCGCTCCCGCGGCGCTAATATCGCGCGCATTTGTGTCGCGACAGT 591
54	479	CTCCCAATGTCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 538
55		
56	592	GCGATTCCTTCCCGCGCGCTCAAGCCCGGCTGCAAGTGGCGGTTTACTGTTCCAGA 651
57	539	GCGATGAGTTCGCCGAGGCCCTCAAGCCCGGCTGCTACTGCGCTTCACTGATTCAGA 598
58		
59	652	AGCGCGAACAACCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
60	599	AGCGCGAACAATCGACTGAGTCTTCTGTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 658

[illegible][illegible]

APPLICANT: M. Nyl, Arja
APPLICANT: Lantto, Raija
APPLICANT: Elvoinio, Minna
APPLICANT: Joutsinki, Vesa
APPLICANT: Paloheimo, Marja
APPLICANT: Suominen, Pirkko
TITLE OF INVENTION: NOVEL CELLULASES, THE GENES ENCODING THEM AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/841,636A
FILING DATE: 30-Apr-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/305,335
FILING DATE: 17-Oct-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/007,926
FILING DATE: 04-Dec-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/020,840
FILING DATE: 28-Jun-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/732,181
FILING DATE: 16-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/F196/00550
FILING DATE: 17-Oct-1996
ATTORNEY/AGENT INFORMATION:
NAME: Timothy J. Shea, Jr.
REGISTRATION NUMBER: 41,306
REFERENCE/DOCKET NUMBER: 1716.0510005/MAC/JUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Melanocarpus albomyces
STRAIN: ALK04237
FEATURE:
NAME/KEY: exon
LOCATION: 33..115
OTHER INFORMATION: /codon_start= 33
OTHER INFORMATION: /product= "20K-cellulase"
FEATURE:
NAME/KEY: exon
LOCATION: 187..435
OTHER INFORMATION: /product= "20K-cellulase"
FEATURE:
NAME/KEY: exon
LOCATION: 506..881
OTHER INFORMATION: /product= "20K-cellulase"
US-08-841-636A-30
Query Match 30.1%; Score 353.6; DB 8; Length 936;

Best Local Similarity 72.8%; Pred. No. 42e-73;
Matches 505; Conservative 0; Mismatches 119; Indels 70; Gaps 1;
QY 145 CGAGTACTGGAAGCTGTCGAAAGCGGTGTCCTTGGCCGGAAGCGCCGTGAGCC 204
Db 183 CGAGTACTGGAAGCTGTCGAAAGCGGTGTCCTTGGCCGGAAGCGCCGTGAGCC 242
QY 205 AACCGGTCTACGCTGATGATGCACTTCCAGGCGCTGTCGATTCGAAATGTCATTCG 264
Db 243 AACCGGTCTACGCTGATGATGCACTTCCAGGCGCTGTCGATTCGAAATGTCATTCG 302
QY 265 GTCGACAGGCGCTGCGGCTTCTACTCTGCGCCACGACGCTTCTGCGCGGATGACGAC 324
Db 303 GTCGACAGGCGGCGCCGCTTCTGTCGCGCCACGACGCTTCTGCGCGGATGACGAC 362
QY 325 ATCTGCGCTTACGCTTTCGCGGACGACATGCGCGCGGATTCGATCTGCTGCTGCT 384
Db 363 ACCTGCGCTTACGCTTTCGCGGACGACATGCGCGCGGATTCGATCTGCTGCTGCT 422
QY 385 GCGCTGCTACGCG----- 398
Db 423 GTCGCTGCTACGCGTGAATGTCGTCGCGCCACGCTGCTGATTCGCGATTCGACGAC 482
QY 399 -----CTCACCTTCACTTCCGCTCCGCTCCGCGGACGAC 434
Db 483 TGAACGAGGACCGGCTGCGGACGCTGACCTTTACATCGGCTCCGTCGCGGACGAC 542
QY 435 ATGATGTCGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 494
Db 543 ATGATGTCGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 602
QY 495 ATGCG 554
Db 603 ATGCG 662
QY 555 GGGCGCTCATATGCGCGCGCATTTGTCGCGCGACGATGCGATTCCTCCCGCGCGCTC 614
Db 663 GGGCGCGCTCATATGCGCGCGCATTTGTCGCGCGACGATGCGATTCCTCCCGCGCGCTC 722
QY 615 AACCGCGCTGACGATGCGCGCGGTTTGAATGCTTCCAGAAAGCGGACGACGACGACG 674
Db 723 AACCGCGCTGACGATGCGCGCGGTTTGAATGCTTCCAGAAAGCGGACGACGACGACG 782
QY 675 TTCGACGATGCGATGCG 734
Db 783 TTCGACGATGCGATGCG 842
QY 735 GACTCCAGCTTCCCGCTTTCACCCCGCCAGCG 768
Db 843 GAGCGCGCTTCCCGCTTTCACAGGCGCCCAAGCG 876
RESULT 10
US-10-007-521-3
Sequence 3, Application US/10007521
Publication No. US20030054539A1
GENERAL INFORMATION:
APPLICANT: Schielein, Martin
Andersen, Lene N.
Laessen, Soren F.
Kauppinen, Markus S.
Lange, Lene
Nielsen, Ruby I.
Ihara, Michiko
Takagi, Shinobu
TITLE OF INVENTION: No. US20030054539A1 Endoglucanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. US20030054539A1 No. US20030054539A1disk of No. US200300C
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America

ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..891
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-007-521-3

Query Match 28.7%; Score 337; DB 15; Length 894;
Best Local Similarity 64.8%; Pred. No. 3,3e-69;
Matches 559; Conservative 0; Mismatches 280; Indels 24; Gaps 3;

QY 108 CTGTGCTGCTCCGCGGCGAGTGGCGAGTCCAGCAAGTACTGCGAGTGGCGAGG 167
DB 43 CTGGCCCTGAGCAAGCTCTCGGCGAGTCCGCGAGCAAGCAAGCTGAGTGGAGTGGCGAGG 102
QY 168 CCGTGGCTGGTGGCGCGGAGGAGCGCGCTGAGCAAGCCGGTCAAGCTGAGTGGCGAGG 227
DB 103 CGAGCTGGCTGGCGCGGAGGAGCGCGCTGAGCAAGCCGGTCAAGCTGAGTGGCGAGG 162
QY 228 AACTTCAGAGCGCTGCTGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 284
DB 163 GAGCAAGCGCTGAGCAAGCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 222
QY 285 TACTCTGGCGCGAGCAAGTCCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 344
DB 223 TACATGCTCTCCGAGAGCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 282
QY 345 GCGAGAGAGTCCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 404
DB 283 GCGAGAGTCCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 342
QY 405 TTACTCTCGGCTCCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 464
DB 343 TTCAAGAGCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 402
QY 465 GAGTGGAGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 524
DB 403 GAGTGGAGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 462
QY 525 GAGTGGAGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 575
DB 463 GCGTGGAGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 522
QY 576 TCGTGGAGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 635
DB 523 CATTCAGAGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 582

QY 636 TTGACTGCTTCAGAGCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 695
DB 583 TTGACTGCTTCAGAGCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 642
QY 696 GCGAGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 755
DB 643 TCGAGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 702
QY 756 ACCCGCGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 815
DB 703 AACCGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 762
QY 816 GCGAGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 875
DB 763 CCGAGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 810
QY 876 GCGAGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 935
DB 811 GCGAGAGTTCGCGAGTTCAGATGTCAGTGGCGAGTGGCGAGG 870
QY 936 GACTGACTGCGAGTTCGCTGTA 958
DB 871 GACTGACTGCGAGTTCGCTGTA 893

RESULT 11

US-10-007-521-5
Sequence 5, Application US/10007521
Publication No. US20030054539A1
GENERAL INFORMATION:
APPLICANT: Schuelein, Martin
Lassen, Soren F.
Kauppinen, Markus S.
Lange, Lene
Nielsen, Ruby I.
Ihara, Michiko
Takagi, Shinobu
TITLE OF INVENTION: No. US20030054539A1 Endoglycanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSER: No. US20030054539A10 No. US20030054539A1disk of No. US200300
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 927 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 1..924
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-007-521-5

Query Match 28.5%; Score 334.8; DB 15; Length 927;

Best Local Similarity 64.9%; Pred. No. 1..1e-68;

Matches 574; Conservative 0; Mismatches 277; Indels 33; Gaps 4;

109 CTGGTCCGCTCCGCGGCAAGTGGAGTGGCGAGTCCAGAGATCTGAGGACTGCTGCAAG 167
43 CTGGCCCTGAGACAGCTCTCGGCGCATCGGCGAGAGACGACCGGATCTGAGGACTGCTGCAAG 102
168 CGTCTGCTCCCTTGGTCCGCGGAGGCGCGCTGACGCAACCGGCTTACCGCGTGCATGCC 227
103 CGAGCTGCGCCCTGCGCGGAGAGGCGCCCTGCTCTCGGCTGAGGCGCTGCGCAAGAAC 162
228 AACTTCAGAGCGCTGCTCGAGCTTCAATGTCAGTCCGCGTGCAGC---GGCGGCTCGGCG 284
163 GACGACCCGCTCAGAGAGGCGGCTCCAGCCGCTCGGCTCGAGCGCGGCGGCGGCGGCGC 222
285 TACTCTGCGCGGCAAGATCTCTGCGCGGTTGAAAGCAATCTCGCTTACCGCTTCCG 344
223 TACATGCTCTCCGAGAGCGCTCGGCGGCTGACGAGCTGCTGTAACGCTGCGGCG 282
345 GCGAGAGCATCGCGCGGCGGCGGCTCCGAACTCTGCTGCTGCGCTGCGCTGCGCTGCA 404
283 GCGCTCAAGCTCGCGCGGAGCTCTCGAGTGGAGTGTGCGCGCTGCTGCTGCAAGTGA 342
405 TTCACTTCGCGCTCCGCTCGCGCGGCAAGATGATGATGATGATGATGATGATGATGATG 464
343 TTCACTTCGCGCGGCGGCTCGCGCGGCAAGATGATGATGATGATGATGATGATGATGATG 402
465 GACCTGGGAAATTAACAGTTCGATATGCGATGCGCGCGGCGGCGGCGGCGGCGGCGG 524
403 GACCTGGGCGAGCAACCACTTGAACCTGCGCATCCCGGCTGCGGCTGCGGCTGCGGCT 462
525 GCGCTGAGCTCGGAGTGGCGGCGCTCCG---CGGCTGCAATCGCGCGGCGGCGGCT 575
463 GCGCTGAGCTCGGAGTGGCGGCGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 522
576 TCGTCCGCGGCGAGTGGCGGATTCCTTCCGCGGCGGCTCAAGCGCGGCTGCGGCGGCG 635
523 CATTCGAGAGAGAGTGGGATCTTCCGAGAGGCGCTCAAGCGCGGCTGCGGCGGCGGCG 582
636 TTGATCTGCTTCCAGAGCGCGGCAACCGCGGCTTCAAGCGGCTGCGGCGGCGGCGGCG 635
583 TTCACTGCTTCCAGAGCGCGGCAACCGCGGCTTCAAGCGGCTGCGGCGGCGGCGGCG 642
696 GCGGAGATCGTGGCGGCGCTCGGCGTGCAGAGCGGAGCTTCCGCGGCTTCCGCGGCT 755
643 TCGAGCTGAGCTCGGAGAGGCGGCTCGGCGGAGAGAGCGGAGAGCTTCCGCGGCTC 702
756 ACCCGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 797
703 CAGATCCCTCCGAGAGAGAGCGGCTCTCGGCTCAAGCGGCTTCAAGCGGAGCGGAGCG 762
798 TCGAGCTGCGGCTGCGGCGGAGAGC---GTCTCCGCGGCGGCGGCGGCGGCGGCGGCG 854
763 TCGAGCTGCGGCTGCGGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 822
855 CAGAGTGGGCTCAGTGGCGGAGTGGCGGCTTCAAGCGGAGTGCAGCACTGCTGCTGCG 914
823 GAGAGTGGGCTCAGTGGCGGAGTGGCGGCTTCAAGCGGAGTGCAGCACTGCTGCTGCG 882
915 ACCAGCTGCGGAGAGTGAAGAGTCACTTCAAGTGCAGTGCCTTCA 958
883 AGCACTTGAAGAGATTAAGTCACTTCAAGTGCAGTGCCTTCA 926

RESULT 12
US-09-735-787-3

Sequence 3, Application US/09735787
Patent No. US20010036910A1

GENERAL INFORMATION:

APPLICANT: Rasmussen, Gretche

Mikkelsen, Jan Moller

Schulein, Martin

Packar, Shankant A.

Hagen, Fred

Endoglucanase Enzyme

TITLE OF INVENTION: A Cellulase Preparation Comprising an

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESS: No. US20010036910A1 No. US20010036910A1disk of No. US20010036910A1

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/735,787

FILING DATE: 13-Dec-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/189,028

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 3469,214-US

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1473 base pairs

TYPE: nucleic acid

STANDARDS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Pseudomonas aeruginosa

STRAIN: DSM 2672

FEATURE:

NAME/KEY: CDS

LOCATION: 97..1224

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-735-787-3

Query Match 26.8%; Score 315.2; DB 9; Length 1473;

Best Local Similarity 68.5%; Pred. No. 5.1e-64;

Matches 466; Conservative 0; Mismatches 208; Indels 6; Gaps 2;

83 CAGAACCTGCGCGGCTGCACTTCTTGGTGGCTCCGCGCGGCAAGTGGCGGCGGCGG 142
108 CACTCTTCCGCGGCGGCGGCTTCTGCGGAGTGGCTTGGAGAGGAGGAGGAGGAGG 167
143 CAGCAGATCTGAGGAGTGGCGGAGGCGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGG 202
168 TACTCATCTGAGGAGTGGCGGAGGCTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGG 227
203 CCAACCGGCTTACGCGGCGGAGGCAATCTTCAAGGCGGCTTGGAGGAGGAGGAGGAG 262
228 CGCCCTGCTTAACTTGTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 287
263 GGGCTGCAAG---CGGCGGCTGCGGCTTACCTGCGGCGGAGAGAGAGAGAGAGAGAG 319

Db	288	CGGTTGAGAGGAGGTGGTGGTCTCTGTTATGCTTGGACCAACATCACTCTCCCTGGGCTGTGAA	347
Oy	320	CGAACATCTCCCTTACAGGCTTCCGCGCGACACAGCATTCGCGCGGATTCGCAATCTCTG	378
Db	348	CGATGAGACTTGGCTACGGATCTTCCTGTCTACCAAGATCTCCGGTGGCTTCGAGGCGACGCTG	407
Oy	380	GTGTGCGGCTCTACGGGCTCACTTCCTTCCTCCGTCGCGCGCGCAAGACATGGT	435
Db	408	GTGCTGTGGCTCTATGGTTTGAACCTTACCACTCGGCGCCCGTCAAGGCGCAAGAGATGAT	467
Oy	440	GGTGCAGTCAACAGCACTGGCGGCGAAGCTGGGAAGTAAACAGTTGGATATGCCATGCG	499
Db	468	GGTCCAGTCCACCAACACTGAGAGTATCTGGGCGAACAACATTCGATCTCATGTATGCG	527
Oy	500	CGGCGCGGCGGTGGACATCTTCAAAGGCTGGAGCTGGAGTTGCGCG---GAGCTCCCGG	556
Db	528	CGGCGGAGGTGGTGGATCTTTCGACGAGGCTGACCTCTGAGTTGCGCAAGGCTTTCGGCGG	588
Oy	557	CGGTCAATACGGCGGCAATTTCGTGGCGGACCAAGTGGCATTCCTTCCCGCGCGCTCAA	616
Db	588	TGCCAGATACGGCGGGTATCTCTCCCGAAGGATGTGATCACTACCCGAGCTTCTCAA	644
Oy	617	GCGCGGCTCGAGTGGCGGTTTGACGTGTCCAAATGGCGCAACCCGACGTTCACTGT	678
Db	648	GGAAGGTTGCACTGAGCATTCGATCTGGTGAAGAGCGCAACCTGATCTTCACTT	707
Oy	677	CCAGCAGGTGCAGTATCCCGCGCGAGATGGTGTGCCCGGCTCGGCTGCAAGGCAACAGCA	736
Db	708	TGACCAAGGTTCAGTATCCCGCAAGGCTCTCTCGACATCAAGATGATCAAGGTGATGACGA	767
Oy	737	CTCCAGGTTCCCGGCTTCCA	756
Db	768	CTCCAGGTTCCCGGCTTCCA	787

RESULT 13
 US-10-138-870-3
 Sequence 3, Application US/10138870
 Publication No. US20030119167A1
 GENERAL INFORMATION:
 APPLICANT: Rasmussen, Grethe
 Mikkelsen, Jan Moller
 Schulein, Martin
 Parkar, Shankant A.
 Hagen, Fred
 TITLE OF INVENTION: A Cellulase Preparation Comprising an
 Endoglucanase Enzyme
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. US20030119167A1c No. US20030119167A1disk of No. US20030119167A1
 STREET: 405 Lexington Avenue, 64th Floor
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/138,870
 FILING DATE: 03-May-2002
 CLASSIFICATION: <UNKNOWN>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/735,787
 FILING DATE: 13-Dec-2000
 APPLICATION NUMBER: 09/189,028
 FILING DATE: <UNKNOWN>
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 3469,214-US

```

1 TELECOMMUNICATION INFORMATION:
2     TELEPHONE: 212-867-0123
3     TELEFAX: 212-878-9655
4 INFORMATION FOR SEQ ID NO: 3:
5     SEQUENCE CHARACTERISTICS:
6         LENGTH: 1473 base pairs
7         TYPE: nucleic acid
8         STANDARDS: single
9         TOPOLOGY: linear
10        MOLECULE TYPE: cDNA
11        HYPOTHETICAL: NO
12        ANTI-SENSE: NO
13        ORIGINAL SOURCE:
14            ORGANISM: Fusarium oxysporum
15            STRAIN: DSM 2672
16        FEATURE:
17            NAME/KEY: CDS
18            LOCATION: 97..1224
19        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
20        DS-10-138-870-3

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 08:03:07 ; Search time 4752 Seconds

(without alignments)
7377.567 Million cell updates/sec

Title: US-10-007-521-11

Perfect score: 1174

Sequence: 1 GAGCAGCAGCCCTCAAGCTG.....AAAAAAAAAAAAAAAAAAAA 1174

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estha:*
2: em_esthm:*
3: em_estim:*
4: em_estov:*
5: em_estpl:*
6: em_estro:*
7: em_estro:*
8: em_hic:*
9: gb_estcl:*
10: gb_estc2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vitl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227.4	19.4	688	13	BU064446 Ref 4 005
2	213	18.1	691	10	B5585661 B5746FSP6
3	207.6	17.7	589	14	CB928003 B5A1_35_G
4	138.2	11.8	668	14	CB408403 IPG004F12

5	137	11.7	667	14	CB407613	CB407613 IPG028F05
6	132.4	11.3	681	14	CB408544	CB408544 IPG007A08
7	130.4	11.1	450	12	BI200729	BI200729 G1F05Fs.r
8	124.2	10.6	622	14	CB408447	CB408447 IPG006G306
9	121.4	10.3	703	14	CD486264	CD486264 CF058_1FL
10	114.4	9.7	444	12	BI190695	BI190695 13910Fs.r
11	111.8	9.5	593	14	CB407636	CB407636 IPG029B01
12	111.6	9.5	630	14	CB407706	CB407706 IPG027C04
13	109.8	9.4	446	14	CB408057	CB408057 IPG013D01
14	109	9.3	534	14	CB408340	CB408340 IPG010A02
15	109	9.3	568	14	CB407804	CB407804 IPG002B05
16	107.4	9.1	426	12	BI187295	BI187295 alh11fs.r
17	106.4	9.0	663	14	CB407573	CB407573 IPG027G07
18	105.8	8.5	582	14	CB407953	CB407953 IPG023B04
19	100.2	8.5	582	14	CB408097	CB408097 IPG013H06
20	99.6	8.5	515	14	CB408037	CB408037 IPG016D03
21	98.2	8.4	433	14	CB408489	CB408489 IPG017E01
22	98.2	8.4	452	14	CB408817	CB408817 IPG017E01
23	95.4	8.1	571	14	CB407840	CB407840 IPG024C11
24	95.4	8.1	673	14	CB408389	CB408389 IPG004B07
25	79.2	6.7	601	14	CB407923	CB407923 IPG021H08
26	78.8	6.7	385	14	CB408025	CB408025 IPG011H04
27	74.4	6.3	447	14	CB407737	CB407737 IPG025C09
28	73.4	6.3	515	13	CB424977	CB424977 BX424977
29	72.6	6.2	909	13	BX31337	BX31337 BX31337
30	72	6.1	562	14	CD682259	CD682259 r148905.y
31	71.8	6.1	1201	13	BX396673	BX396673 BX396673
32	71.6	6.1	538	13	BX404314	BX404314 BX404314
33	70.4	6.0	191	13	BX477475	BX477475 DKZ2P686K
34	70.4	6.0	193	13	BX509937	BX509937 DKZ2P686K
35	70.2	6.0	168	13	B0152857	B0152857 NF025H041
36	70	6.0	889	14	CF883951	CF883951 LK1C036Xh
37	69.8	5.9	297	14	CB408884	CB408884 IPG01BD09
38	69.6	5.9	203	9	AT620089	AT620089 ty48d08.x
39	69.4	5.9	520	14	CD683022	CD683022 r179e08.y
40	69.4	5.9	974	13	BX339244	BX339244 BX339244
41	69.2	5.9	233	14	CF335596	CF335596 JMT - 05-F
42	68.8	5.9	493	12	BG656882	BG656882 df25908.y
43	68.4	5.8	804	13	B0414074	B0414074 60366048
44	68.4	5.8	441	12	BM573394	BM573394 fx34F05.y
45	68.4	5.8	596	13	B0186959	B0186959 UT-E-BU1-

ALIGNMENTS

RESULT 1
BU064446
LOCUS
DEFINITION
Fgr 4 005_T3 Nitrogen-starved mycelia *Gibberella zeae* cDNA, mRNA
VERSION
BU064446.1 GI:22504735
KEYWORDS
EST.
SOURCE
Gibberella zeae
ORGANISM
Gibberella zeae
REFERENCE
1 (bases 1 to 688)
AUTHORS
Trail, F., Xu, C.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.
TITL
Analysis of expressed sequence tags from *Gibberella zeae* (anamorph
Fusarium graminearum)
JOURNAL
Fungal Genet. Biol. 38 (2), 187-197 (2003)
MEDLINE
22508120
PUBMED
12620255
COMMENT
Contact: Frances Trail
Department of Plant Biology
Michigan State University
East Lansing, MI 48824, USA
Tel: 517 432 2939
Fax: 517 353 1926
Email: trail@msu.edu
Plate: 4 row: 0 column: 05.

FEATURES

Location/Qualifiers

1. 688

/organism="Gibberella zeae"

/mol_type="mRNA"

/strain="NRRL 31084"

/db_xref="taxon:5518"

/clone_lib="Nitrogen-starved mycelia"

/note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16; Mismatches 191; Indels 6; Gaps 2;

Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

FEATURES

Location/Qualifiers

1. 691

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Sumai3"

/db_xref="taxon:4565"

/clone="EST#6SP6_D02_d2_014"

/tissue_type="Spike"

/dev_stage="Adult plant"

/lab_host="E. coli JM109"

/clone_lib="K5U wheat Fusarium graminearum infected spike"

/note="Vector: pGEM-T easy; Site 1: SacII; Site 2: SpeI; Plants were grown in the greenhouse. Spikes were sprayed with Fusarium graminearum (at what stage). Total RNA, and poly(A) RNA were prepared from infected spikes. cDNA was prepared using the SmartTM PCR cDNA synthesis kit from Clontech. cDNA was cloned into the pGEM-T easy vector from Promega."

ORIGIN

Query Match

18.1%; Score 213; DB 10; Length 691;

Best Local Similarity 64.9%; Pred. No. 6.7e-15; Mismatches 190; Indels 7; Gaps 3;

Matches 364; Conservative 0; Mismatches 190; Indels 7; Gaps 3;

FEATURES

Location/Qualifiers

1. 691

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Sumai3"

/db_xref="taxon:4565"

/clone="EST#6SP6_D02_d2_014"

/tissue_type="Spike"

/dev_stage="Adult plant"

/lab_host="E. coli JM109"

/clone_lib="K5U wheat Fusarium graminearum infected spike"

/note="Vector: pGEM-T easy; Site 1: SacII; Site 2: SpeI; Plants were grown in the greenhouse. Spikes were sprayed with Fusarium graminearum (at what stage). Total RNA, and poly(A) RNA were prepared from infected spikes. cDNA was prepared using the SmartTM PCR cDNA synthesis kit from Clontech. cDNA was cloned into the pGEM-T easy vector from Promega."

FEATURES

Location/Qualifiers

1. 688

/organism="Gibberella zeae"

/mol_type="mRNA"

/strain="NRRL 31084"

/db_xref="taxon:5518"

/clone_lib="Nitrogen-starved mycelia"

/note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16; Mismatches 191; Indels 6; Gaps 2;

Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

FEATURES

Location/Qualifiers

1. 688

/organism="Gibberella zeae"

/mol_type="mRNA"

/strain="NRRL 31084"

/db_xref="taxon:5518"

/clone_lib="Nitrogen-starved mycelia"

/note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16; Mismatches 191; Indels 6; Gaps 2;

Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16; Mismatches 191; Indels 6; Gaps 2;

Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

JOURNAL

spike cDNA library

Unpublished (2000)

Contact: John Fellers

US Department of Agriculture, Agriculture Research Service, Plant

Science and Entomology Unit

Dept. of Plant Pathology, 4006 Throckmorton Hall, Kansas State

University, Manhattan, KS 66506, USA

Tel: 785-532-2367

Fax: 785-532-6167

Email: jpfellers@ksu.edu

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: SP6.

FEATURES

Location/Qualifiers

1. 691

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Sumai3"

/db_xref="taxon:4565"

/clone="EST#6SP6_D02_d2_014"

/tissue_type="Spike"

/dev_stage="Adult plant"

/lab_host="E. coli JM109"

/clone_lib="K5U wheat Fusarium graminearum infected spike"

/note="Vector: pGEM-T easy; Site 1: SacII; Site 2: SpeI; Plants were grown in the greenhouse. Spikes were sprayed with Fusarium graminearum (at what stage). Total RNA, and poly(A) RNA were prepared from infected spikes. cDNA was prepared using the SmartTM PCR cDNA synthesis kit from Clontech. cDNA was cloned into the pGEM-T easy vector from Promega."

ORIGIN

Query Match

18.1%; Score 213; DB 10; Length 691;

Best Local Similarity 64.9%; Pred. No. 6.7e-15; Mismatches 190; Indels 7; Gaps 3;

Matches 364; Conservative 0; Mismatches 190; Indels 7; Gaps 3;

FEATURES

Location/Qualifiers

1. 691

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Sumai3"

/db_xref="taxon:4565"

/clone="EST#6SP6_D02_d2_014"

/tissue_type="Spike"

/dev_stage="Adult plant"

/lab_host="E. coli JM109"

/clone_lib="K5U wheat Fusarium graminearum infected spike"

/note="Vector: pGEM-T easy; Site 1: SacII; Site 2: SpeI; Plants were grown in the greenhouse. Spikes were sprayed with Fusarium graminearum (at what stage). Total RNA, and poly(A) RNA were prepared from infected spikes. cDNA was prepared using the SmartTM PCR cDNA synthesis kit from Clontech. cDNA was cloned into the pGEM-T easy vector from Promega."

FEATURES

Location/Qualifiers

1. 688

/organism="Gibberella zeae"

/mol_type="mRNA"

/strain="NRRL 31084"

/db_xref="taxon:5518"

/clone_lib="Nitrogen-starved mycelia"

/note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16; Mismatches 191; Indels 6; Gaps 2;

Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

FEATURES

Location/Qualifiers

1. 688

/organism="Gibberella zeae"

/mol_type="mRNA"

/strain="NRRL 31084"

/db_xref="taxon:5518"

/clone_lib="Nitrogen-starved mycelia"

/note="Vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16; Mismatches 191; Indels 6; Gaps 2;

Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16; Mismatches 191; Indels 6; Gaps 2;

Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

QY 623 CTGCCAGTGGCGGTTTACTG 643
DB 670 TTGCCACTGGGATTTGACTG 690

RESULT 3
CB928003 589 bp mRNA linear EST 28-APR-2003
LOCUS AB01_35.G07.g1.A012 Abscissic acid-treated seedlings Sorghum bicolor
DEFINITION cDNA clone AB01_35.G07.A012 5', mRNA sequence.
ACCESSION CB928003
VERSION CB928003.1 GI:30164274
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 589)
Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Buchanan, C.D.,
Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S.,
Eastman, A. and Pratt, L.H.
An EST database from Sorghum: AB01-treated seedlings
Unpublished (2003)
COMMENT Contact: Cordomier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 593 0210
Email: emppratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTCTGCTCTAAAGCTGCG).
FEATURES
source
1..589
location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="IS3620C"
/db_xref="taxon:4558"
/clone="AB01_35.G07.A012"
/lab_host="DH10B-T1 phage-resistant E. coli"
/note="Vector: pME18S-FU3; Site_1: XhoI; Site_2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. After 12 days, medium was
supplemented with 1 mM abscissic acid (ABA), while leaves
were misted with a solution of 1 mM ABA. Roots and leaves
were harvested after 3, 6, 12, and 24 hr and material from
all time points was combined prior to RNA isolation.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FU3 vector (5'-prime
DraIII site is CACTGTGTG, 3'-prime DraIII site is
CACCATGTG). XhoI excises the cDNA insert."

ORIGIN
Query Match 17.7%; Score 207.6; DB 14; Length 589;
Best local Similarity 66.5%; Pred. No. 2.8e-14;
Matches 310; Conservative 0; Mismatches 153; Indels 3; Gaps 1;
QY 83 CACAACCTGGCGGCTGCTCTGCTGCTCGGCGGCGAGTGGCAGTGGCCAGTGC 142
DB 119 CACTCTTCTGCGCCGCGGCGGCTCTCCCGTAGAGTGTCTCTGAGACGCGCACTC 178
QY 143 CACGAGATATCTGGACTGTCTGCAAGCCGTCTGGCGCTTGGCCGCGGAAAGCCCGCTCAG 202

DB 179 TACTCGATACGTGGGATGTGTCGAAGCCTTCTTGTCTTGAGAGTGGAAGGCTGTGNC 238
QY 203 CCAACGGGTCTACCGGTGCGATGCGCAACTTCCAGCGGCTGTGCGACTTCAATGTCAATC 262
DB 239 TGCCCTGTGCTTGTGACTTGTGCAAGAACGACCAACCCATCTCCACACCAAGCGCCNCA 298
QY 263 GGGCTCGCAA---CGGCGGTGCGCTACTCTGCGCGCGACGACACTCCCTGGGCGGTGAA 319
DB 299 CGATGAGAGAGGCGGTGTTGCTTGTGCTTGTGCTTGTGCAATTAATCTCTCTGGGCTGTGA 358
QY 320 CGACATATCTGCGCTTACCGCTTGTGCGCGAGACATGCGCGCGGCTGCGAATCTCTGCG 379
DB 359 CGATGAGACTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 418
QY 380 GTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 439
DB 419 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 478
QY 440 GGTGCAATGTCAGAGACATGCGCGCGCGACCTGCGGAAAGTAAACAGTTGATATGCGCATGCC 499
DB 479 GGTGCAATGTCAGAGACATGCGCGCGCGACCTGCGGAAAGTAAACAGTTGATATGCGCATGCC 538
QY 500 CGCGCGCGCGGCTGCGATCTTCAACGCGCTGCGAGCTGCGAGTTTCGCGC 545
DB 539 CGCGCGCGGCTGCGATCTTCAACGCGCTGCGAGCTGCGAGTTTCGCGC 584

RESULT 4
CB408403 668 bp mRNA linear EST 24-OCT-2003
LOCUS CB408403
DEFINITION clone IPG04F12 249708 JH III-treated male I. pini midguts Ips pini cDNA
ACCESSION CB408403
VERSION CB408403.1 GI:37952539
KEYWORDS EST.
SOURCE Ips pini (North American pine engraver)
ORGANISM Ips pini
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Phytophaga; Scolytidae; Ips.
REFERENCE 1 (bases 1 to 668)
Bigenbee, A.L., Keeling, C.I., Young, S. and Tittiger, C.
Comparison of gene representation in midguts from two phytophagous
insects, Bombyx mori and Ips pini, using expressed sequence tags
Gene 316, 127-136 (2003)
JOURNAL MEDLINE
PUBMED 22925480
COMMENT Contact: Tittiger C
Biochemistry Department
University of Nevada, Reno
Howard Medical Sciences Rm. 150, Reno, NV 89557, USA
Tel: 1-775-784-6480
Fax: 1-775-784-1419
Email: crit@unr.edu
PCR PRIMERS
FORWARD: T3 21mer
BACKWARD: T7 21mer (backward)
Plate: 004 row: F column: 12
Seq primer: T3 21mer
High quality sequence stop: 668.
location/Qualifiers
1..668
source
/organism="Ips pini"
/mol_type="mRNA"
/db_xref="taxon:102803"
/clone="IPG04F12"
/sex="male"
/tissue_type="midgut"
/dev_stage="adult"
/note="Vector: lambda Uni-Zap XR, Bluescript SK-; As per
Stratagene's pBluescript II XR library construction kit"

ORIGIN

LOCUS	CB408544	661 bp	mRNA	linear	EST 24-OCT-2003
DEFINITION	IPB007A08.249626 UH III-treated male I. pini midguts Ipe pini cDNA clone IPB007A08 5, mRNA sequence.				
ACCESSION	CB408544				
VERSION	CB408544.1	GI:37952680			
KEYWORDS	EST.				
SOURCE	Ipe pini (North American pine engraver)				
ORGANISM	Ipe pini				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformata; Phyllophaga; Scolytidae; Ips.				
AUTHORS	1 (bases 1 to 661)				
TITLE	Eisenberg A.L., Keeling C.I., Young S. and Tittiger C.				
JOURNAL	Comparison of gene representation in midguts from two phytophagous insects, <i>Bombix mori</i> and <i>Ips pini</i> , using expressed sequence tags				
MEDLINE	Gene 316, 127-136 (2003)				
PUBMED	22925480				
COMMENT	14563559				
	Contact: Tittiger C				
	Biochemistry Department				
	University of Nevada, Reno				
	Howard Medical Sciences Rm. 150, Reno, NV 89557, USA				
	Tel: 1-775-784-6480				
	Fax: 1-775-784-1419				
	Email: ctt@unr.edu				
	PCR Primers				
	FORWARD: T3 21mer				
	BACKWARD: T7 21mer (backward)				
	Plate: 007 row: A column: 08				
	Seq primer: T3 21mer				
	High quality sequence scrop: 661.				
FEATURES	Location/Qualifiers				
source	1..661				
	/organism="Ips pini"				
	/mol type="mRNA"				
	/db xref="taxon:102303"				
	/clone="IPB007A08"				
	/sex="male"				
	/tissue type="midgut"				
	/dev stage="adult"				
	/clone lib="UH III-treated male I. pini midguts"				
	/note="Vector: lambda Uni-Zap XR, Bluescript SK-; As per Stratagene's pBluescript II XR library construction kit"				
ORIGIN					
Query Match	11.3%	Score 132.4	DB 14	Length 661	
Best local Similarity	55.1%	Pred. No. 6.3e-06			
Matches 350	Conservative 0	Mismatches 267	Indels 10	Gaps 4	
Dp	99	GCAATTCCTCTGTGTCGTCGCGCGCGACGAGGAGGACAGTCCACAGAGATCTGGGAC	158		
Dp	32	GCTGTGTGTGTGTCACGCTGCGCTCGTGTGATATGATCTACACCCGTTACTGGGAC	91		
Dp	159	TGCTGCAAGCCGTCGTCGCGCTTGACCCGAGAGCGCGCCGTACGCCAATCGGCTTACGCG	218		
Dp	92	TGTTGCAACATCTCTGCTGTGAGAAAGAAATTTGAGAGACCTGAACTGTGTATTC	151		
Dp	219	TGCGATGCCAATCTTCAGAGCGCTGTTCGCACTTCAATGTCTGATGCGGCTGCAAC	275		
Dp	152	TGGCGTAAAGATGTGTTCACCGCTTGAAACCCCTGTTCACATCTGGTGTGATTAAG	211		
Dp	276	GAGCTCGGCGCTATCTCTGCGCGACACGACCTCCCTGGGCGGTGAACGACATCTGCGCTAC	335		
Dp	212	GATTTCTGCTTACGTCGTGCAACAACCTTCCACCTTGGGCTGTCAACGACACTGCGCTAC	271		
Dp	336	GAGCTTCGCGACAGACAGATCGCGCGGAGTCCGAAATCTGTGTGTGCTGCGCTCTAC	395		
Dp	272	GGATTTGTGCTGCTCTTCACTGGTGGTGTGAACAACAGCTACTGTGGTATTTGTTTG	311		
Dp	336	GCGTCACTTCACTTCCGATCCCGTCCGCGGCAAGACAGATGTGTGACATGCAACGAGC	455		

	Db	332	AATTGACCTTCA---CCAAATTAAGTGCGAAGAAGCCTTCATGTCGAAAACATCAAC	388
Oy	456	ACTGGCGCGACCTCGGGAAGTAACAGTTCCATATCCCATATGCCGCGCGCGCTGGGC	515	
Db	389	ACCGGTGGTAGCTTAGAGATCCAAACAATTCGACATTGAATCCCCGAGGTGGTGGT	448	
Oy	516	ATCTTCAAC---GGCTGACAGCTCGCAGTTCCGCGGCGCTCCCGGAGCT-----CAA	563	
Db	449	ATTCTAACCTTGGGTGTTCAAACCATGGAGTCTCTTCTGCGGTTGGGGTCAACA	508	
Oy	564	TACGCGGCATTTCGTGCGCGACAGATGCGATTCTTCCC CGCGCGCTCAAGCCCGC	623	
Db	509	TACGGGAGAGTTTCCTGAGACTGTGATGGATGACTCTCTCCCTGAGAACTCATCTGSA	568	
Oy	624	TTCGAGTGGCGGTTTACATCTGTTCCAGAAAGCCGACAAACCCAGTACGTTCCAGCA	683	
Db	569	TTCGACTTCGATTCGACGCGGTACGAAATAATGTGACATCCCAAGTGAATTTGAACA	628	
Oy	684	GTCGAGTGCACCGCGCAGATGCGTTGCCCGCTCCG	718	
Db	629	ATTGACTGTCCGAGAGACTCACAGTCTACCG	663	
RESULT 7				
BI200729		450 bp	mRNA	linear
LOCUS		o1f05fs.r1	Fusarium sporotrichioides Tri 10 overexpressed cDNA	
DEFINITION		library Fusarium sporotrichioides cDNA clone o1f05fs 5', mRNA		
SEQUENCE				
ACCESSION	BI200729	GI:14666701		
VERSION	BI200729			
KEYWORDS				
SOURCE				
ORGANISM			Fusarium sporotrichioides	
REFERENCE			Fusarium sporotrichioides	
AUTHORS			Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;	
JOURNAL			Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.	
COMMENT			1. (bases 1 to 450)	
			Ren,Q., Tag,A., Pehlouw,A., Lai,H., Kupfer,C., Peterson,A.,	
			Beremand,M. and Roe,B.	
			Analysis of a Fusarium sporotrichioides EST database	
			Unpublished (2001)	
			Contact: Bruce A. Roe, University of Oklahoma, broeou.edu	
			Department of Chemistry and Biochemistry	
			Advanced Center for Genome Technology, University of Oklahoma	
			620 Parrington Oval, Norman, OK 73019, USA	
			Tel: 405 325 4912	
			Fax: 405 325 7762	
			Email: broeou.edu	
			Contact Dr. Marian Beremand regarding clone availability included	
			is the best homolog from a blastx search of Genbank nr 04-09-01	
			K33 5e-66 gi 1170140 sp P45699 PUTATIVE ENDODERMAL TYPE	
			EXPECURSOR (EN	
			Seq primer: T3	
			High quality sequence stop: 440.	
FEATURES				
source			Location/Qualifiers	
			1..450	
			/organism="Fusarium sporotrichioides"	
			/mol_type="mRNA"	
			/strain="Tri 10"	
			/db_xref="taxon:5514"	
			/clone="o1f05fs"	
			/clone_1fb="Fusarium sporotrichioides Tri 10 overexpressed	
			cDNA library"	
			/note="Vector: pBluescript SK-. Site 1: EcoRI, Site 2:	
			XhoI; 5' end of cDNA cloned into EcoRI site of	
			pBluescript; 3' end of cDNA cloned into XhoI site of	
			pBluescript"	
ORIGIN				
Query Match		11.1%	Score 130.4; DB 12; Length 450;	
Best local Similarity		62.4%;	Pred. No. 1.3e-05;	
Matches 222; Conservative		0; Mismatches 131; Indels	3; Gaps	1

